

Enjoy this site? 

Gift the author a WordPress.com plan.

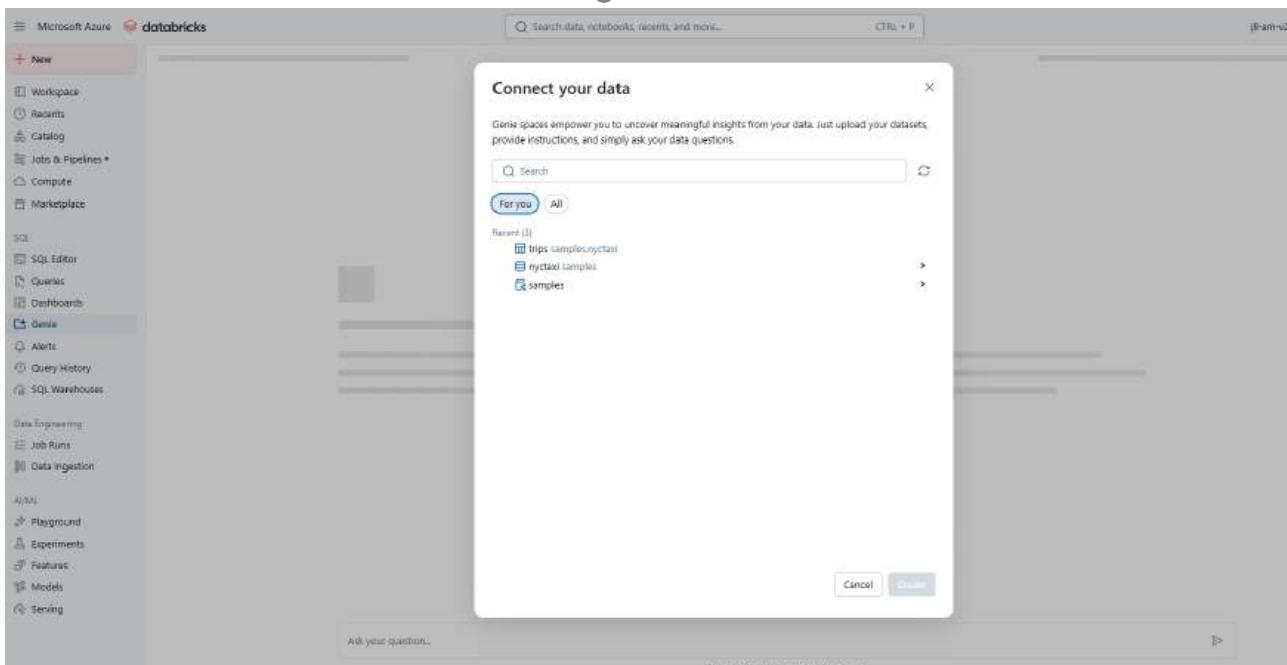
Gift

## DATABRICKS GENIE – GETTING STARTED

Hi all,

We recently migrated at work to databricks and moved away from transformation tools like Alteryx. Part of that learning journey is knowing what parts of databricks we can best utilise. The team have done fantastic work understanding everything from the unity catalog, through to data ingestion and runs, compute and notebook best practice, but of course databricks genie is the ‘glam’ bit.

Today will be a short walk through an existing dataset on databricks genie.



First of all looking at connecting some example data. & we can ask it a generic first question. You have the option to add multiple data sources in here.

## New Space ★

What tables are there and how are they connected? Give me a short summary.

There is one table available: `samples.nyctaxi.trips`. This table contains data about NYC taxi trips, including pickup and dropoff times, trip distance, fare amount, and pickup/dropoff zip codes. Since there is only one table, there are no connections or relationships to other tables.

Was this correct?

We can then ask it something that's probably not that hard a question, but I don't give it any context as to whether the `trip_distance` is a mile or km metrics. Let's see what it comes out with when asked the average cost of a trip per mile.

New Space ★  
relationships to other tables.

Was this correct?

what is the average cost of a trip per mile?

← Analysis complete → trips:

The user wants to find out the average cost of a taxi trip per mile, calculated by dividing the total fare by the distance of each trip, while excluding any trips with zero or no distance.

1. trip\_distance > 0

1 row Add as instruction Hide code

```
1: SELECT
2: AVG(trip_divide("trips"."fare_amount", "trips"."trip_distance")) AS avg_cost_per_mile
3: FROM
4: samples.nyctaxi."trips"
5: WHERE
6: "trips"."trip_distance" IS NOT NULL
7: AND "trips"."trip_distance" > 0
```

Run

1.2 avg\_cost\_per\_mile

6.02

↪ What is the total fare amount for all trips?

↪ What is the average trip distance?

↪ How many trips have a fare amount?

Ask your question...

Always review the accuracy of responses.

The good things is we can ask it to display the code. Taking the average fare amount, and average distance, then dividing them for all values above zero or not null.

## New Space ★

what is the average cost of a trip per kilometre?

Analysis complete trips

The user wants to find the average cost of a taxi trip per kilometer, calculated by converting the trip distance from miles to kilometers and averaging the fare amount accordingly.

1.2 trip\_distance: > 0

1 row Add as instruction Hide code

```
1 SELECT
2   AVG(try_divide(`trips`.`fare_amount`, (`trips`.`trip_distance` * 1.60934))) AS avg_cost_per_km
3 FROM
4   `samples`.`nyctaxi`.`trips`
5 WHERE
6   `trips`.`trip_distance` IS NOT NULL
7   AND `trips`.`trip_distance` > 0
```

Run

	1.2 avg_cost_per_km
1	3.74

Here is that in kilometres.

I could then tell Genie that my original data is in kilometres not miles, and if I re-ask it, it would give me that initial 6.02 value we had previously.

The screenshot shows the Microsoft Azure Databricks interface. On the left, the sidebar includes sections for Workspace, Recent, Catalog, Jobs & Pipelines, Compute, Marketplace, Genie, Alerts, Query History, and SQL Warehouses. Under Data Engineering, there are Add Runtimes and Data Ingestion. Under AI/ML, there are Playground, Experiments, Features, Model, and Serving. The main area shows a new space named 'New Space'. Inside, there is a table with one row and two columns, labeled '1.2 avg\_cost\_per\_km' and '3.74'. Below this table is a 'trip\_details' panel. The panel has a 'Description' section stating 'The distance captured here is in kilometers, not miles. Make sure the calculation reflects that and the conversion between kilometers and miles. To convert kilometers to miles, multiply the distance value in kilometers by 0.621371.' It also has a 'Columns (0)' section listing 'trip\_pickup\_datetime', 'trip\_dropoff\_datetime', 'trip\_distance', 'fare\_amount', 'pickup��', and 'dropoff��'. Each column entry includes a checkbox for 'No description' and a link for 'Example values'. At the bottom of the panel, there are 'Cancel' and 'Save' buttons, and a note 'Is this correct? (1) Yes (2) Fix it (3) Request review...'. Below the panel, there are three questions: 'What is the total fare amount for all trips?', 'What is the average trip distance?', and 'How many trips were taken in total?'. A text input field 'Ask your question...' is at the bottom, and a note 'Always review the accuracy of responses.' is at the very bottom.

Only a small change but important context for the calculation. Next I want to give some general instructions a go. Let's see if this helps.

[Data](#) [Instructions](#) [Settings](#)

X

[Text](#)[Joins](#)[SQL Queries](#)

## General Instructions

Add general instructions on how you want Genie to behave.

Dataset-specific business information:

Peak hours are defined as 7-9 AM and 5-7 PM on weekdays

Long trips are rides over 10 miles or 45 minutes duration

Short trips are rides under 1 mile or 5 minutes duration

Financial calculations:

Tips above 25% of fare should be flagged as potential data anomalies

When calculating average fare per km, exclude trips with \$0 fare or distance

Data quality rules:

Exclude trips with pickup/dropoff at the same location (potential data errors)

Remove records where pickup time is after dropoff time

Formatting preferences:

Always show monetary values with \$ symbol and 2 decimal places

Round trip distances to 1 decimal place

Display percentages with % symbol and 1 decimal place

Format duration as hours:minutes (e.g., 1:23 for 1 hour 23 minutes)

[Save](#)

& we can see that in action when we ask genie to show me the percentage of trips that are classified as long trips.

New Space ★

Show me the percentage of trips that are classified as long trips by month

Analysis complete: trips

The user wants to see the percentage of taxi trips that are considered long trips, broken down by month.

2 rows Add as instruction Hide code

```
28 month,
29 ROUND(
30    trip_startdt
31    * 100.0
32    + sum(
33        CASE
34            WHEN
35                trip_stats.trip_distance > 10
36                OR trip_stats.duration_min > 45
37            THEN
38                1
39            ELSE 0
40        END
41    )
42    / COUNT(*)
43 ) * 1
44 )
```

Run

#	year	month	%_pct_long_trips
1	2016	1	3.8%
2	2016	2	53%

Data Instructions Settings

Text Joins SQL Queries

### General Instructions

Add general instructions on how you want Genie to behave.

Dataset-specific business information:

- Peak hours are defined as 7:00 AM - 7:00 PM on weekdays
- Long trips are rides over 10 km or 45 minutes duration
- Short trips are rides under 1 km or 5 minutes duration

Financial calculations:

- Tip above 25% of fare should be flagged as potential data anomalies
- When calculating average fare per km, exclude trips with 0 fare or distance

Data quality rules:

- Exclude trips with pickup/dropoff at the same location (potential data errors)
- Remove records where pickup time is after dropoff time

Formatting preferences:

- Always show monetary values with \$ symbol and 2 decimal places
- Round trip distances to 1 decimal place

Display percentages with % symbol and 1 decimal place

Format duration as hours:minutes (e.g., 1:23 for 1 hour 23 minutes)

We could also ask it a question around our data quality. (Yes i'm aware someone can be picked up and dropped off at the same location)

New Space \*

What percentage of our dataset consists of valid trips after applying data quality filters

Analysis complete 3 trips

The user wants to know the percentage of trips in the dataset that are considered valid after applying specific data quality filters, which include ensuring that the pickup time is before or equal to the dropoff time and that the pickup and dropoff locations are different.

Time ↑ Add as instruction Hide code

```
2 SELECT
3     COUNT(*) AS total_count
4     FROM
5     samples..`nyctaxi1`..trips
6 );
7 valid AS (
8     SELECT
9         COUNT(*) AS valid_count
10        FROM
11        samples..`nyctaxi1`..trips
12    WHERE
13        `trips`.`tpep_pickup_datetime` <= `trips`.`tpep_dropoff_datetime`
14        AND `trips`.`pickup_zip` != `trips`.`dropoff_zip`
15 );
16 SELECT
17     ROUND(try_divide(100.0 * valid.valid_count, total.total_count), 3) || '*' AS act_valid_trips,
18     PERIOD,
19     total,
20     valid
```

Run

Act valid trips: 99.3%

Data Instructions Settings

Text Joins SQL Queries

**General Instructions**

Add general instructions on how you want Genie to behave:

**Dataset-specific business information:**  
Peak hours are defined as 7-9 AM and 4-7 PM on weekdays  
Long trips are rides over 10 km or 45 minutes duration  
Short trips are rides under 1 km or 3 minutes duration

**Financial calculations:**  
Tips above 25% of fare should be flagged as potential data anomalies  
When calculating average fare per km, exclude trips with 0 fare or distance

**Data quality rules:**  
Exclude trips with pickup/dropoff at the same location (potential data errors)  
Remove records where pickup time is after dropoff time

**Formatting preferences:**  
Always show monetary values with \$ symbol and 2 decimal places  
Round trip distances to 1 decimal place

Display percentages with % symbol and 1 decimal place  
Format duration as hours:minutes (e.g., 1:23 for 1 hour 23 minutes)

# Giving feedback

Each response prompts the user to answer Is this correct?. Users can respond in one of the following ways:

- **Yes:** Confirms the response appears accurate.
  - **Fix it:** Flags the response as incorrect. Users can select from common issues or enter their own explanation. They can then:
    - Click **Submit and try again** to regenerate the response using the provided feedback.
    - Click **Submit** to send the feedback without regenerating the response.
  - **Request review:** Flags the response for manual review. Users can add an optional comment to give you additional context.

As an editor, you can view feedback and flagged responses in the Genie interface.

Overall, I like where genie is headed. I was a little nervous to start with how it would write the code calculations, but the more information you

give it, the more complexities you break down and the response feedback you feed it – the better the outcomes.

At the moment this probably helps a bunch with our teams quality assurance, quick analysis – less so stakeholder facing and intuitive for non data folk.

Probably goes to show the importance of semantic layers in an AI-BI built products and the detailed context needed for good quality reporting.

You can read more about the set up [here](#).

**This video** from Databricks is pretty useful from 17 mins in.

LOGGING OFF,  
CJ

## FUTURE OF AI IN FOOTBALL SCOUTING – RICARDO HEREDIA JOYA

Hi all,

I am joined this week by Ricardo. Ricardo really impressed me with his post match reports. I had the pleasure of chatting to him on a video call last week to learn about his personal data science projects, the role AI will have in football and the opportunity for personal growth.

You can find similar concepts and themes below in our conversation below. Be sure to check out some of Ricardo's main sites:

**Substack**  
**x (Twitter)**  
**Github**



Ricardo Heredia Joya 1st

Data Scientist • AI/ML | Generative AI | Football Analytics | Medical Doctor | Creating data-centric, reliable, and optimized AI systems.

Tres Cantos, Community of Madrid, Spain · [Contact info](#)

[Football Analysis X Profile](#)

500+ connections



Sigma Cognition



Universidad de Carabobo

CJ: Ricardo, thank you for giving up time to share more on your journey with soccer analytics. What was your entry point into the world of data?

It looks like you have had a background in sports health first?

R: Thank you for the opportunity, CJ, I appreciate it!

My background is from medicine, I graduated as a medical doctor in Venezuela and moved out to Spain to keep doing postgraduate studies here. While I was studying for the Spanish medical bar I had a temporary job as a sports physician.

At the same time, I stumbled upon DataCamp Python course and decided to learn it cause it felt like a challenge and I was also yearning for learning a new digital skill.

Got immediately hooked to that course and found out you could apply that to data analytics in sports, specifically in football, my favourite one. So, I made the decision to pivot towards data analysis and going all in.

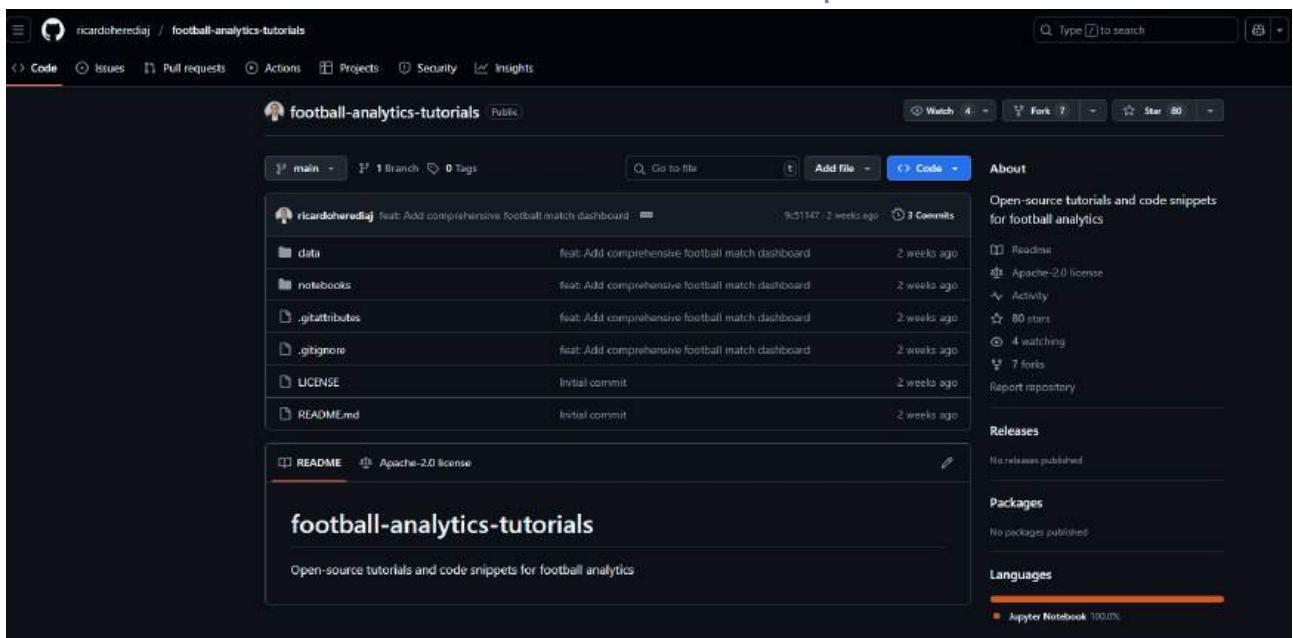
Then, a couple months after I finished a MOOC on Data Analytics with R, I contacted a data scientist that was working with football analytics on LinkedIn. I asked him for advice on where to find free data to start analyzing and to my surprise, he invited me to be part of a data science consultancy group remotely that collaborated with several professional clubs in the US and Europe. I said yes, and it became my first real world experience and entry point working with data in general.

CJ: You often write on **substack** on where to access different resources for football data. Where should someone start if brand new?

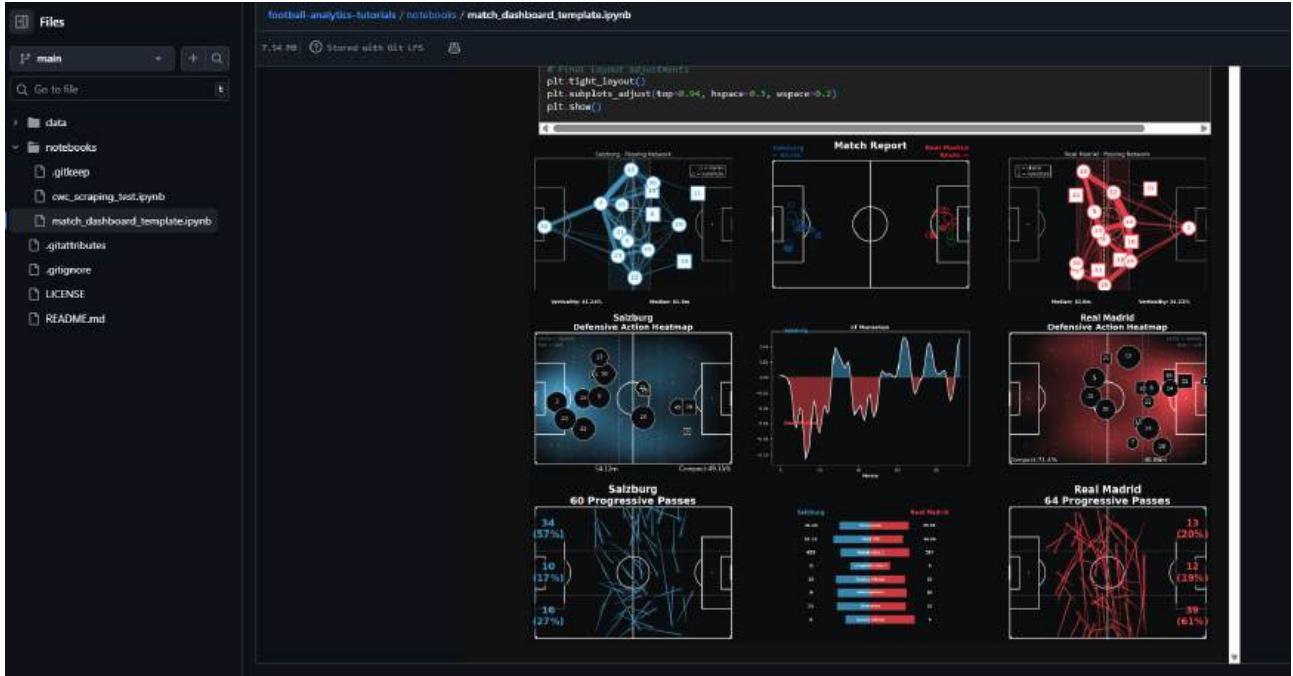
R: As a newbie I didn't have a clue on where to start or where to look for.

That's why I wrote a post exactly about that, where I gathered the resources I used 3 years ago to start playing around with free data, what analyses you can make, where to get both the technical and domain knowledge to make the most out of football data. You'll find free data sources, YouTube channels, online courses and many resources to start.

CJ: The thing I loved most coming across your content was your **Football Post Match report**.



For those that are just getting into coding and soccer analytics. Could you talk a little about what the code does within the repo?



R: The code focuses on 4 things using Python:

1. Obtains free football event data from two popular webs like WhoScored (Opta data with coordinates) and Fotmob (for shots data). It obtains the data through a technique called web scraping, that if done responsibly can help you get the data you need to do any kind of analysis. The **scraping code** is at the beginning of this notebook.
2. Cleaning the scraped data and organizing it for analysis purposes, involving variable renaming, dealing with nulls, and storing the data in CSV and JSON format.
3. Calculations & Plotting: once the data is organized, and you have an idea of what you want to plot (Attacking and Defensive events like shots, progressive passes, the defensive block height, etc) there comes the logic that computes these variables and adapts them to the pitch dimensions. This is important because it helps you set the coordinates and the aesthetic (colours) of the pitch. Then I use Python functions to build each plot and convert them into subplots that would be part of the final match dashboard.
4. Match dashboard: all the subplots are grouped into the same layout giving you the overview of the match with the specific plots I decided to include, in this case only 6 cause I thought it would look less crowded.

A short summary: extracting data -> cleaning and organizing it -> calculations & plotting -> final match dashboard.

When doing projects like this it helps break each part in smaller bits and build upon them.

CJ: As someone that is familiar with matplotlib, when creating a dashboard template like this one. What are some of the easier things to do in terms of design and styling, and what parts of the dashboard take greater thought and consideration?

R: In terms of design and styling the easiest part is looking at different pitch examples to inspire and create yours. For instance, you can go the mplsoccer web (<https://mplsoccer.readthedocs.io/en/latest/>) and look at

the examples. You can also see in social media some of the work other people share and inspire from it to adapt it to your own style (always giving credit to your source of inspiration).

For the dashboard, the hard part was the data scraping process, because you need to look at the layout of the webpage, where the tabs to get the data and variables you need are, what the output returns, etc.

Thankfully I didn't start from zero, got the opportunity to do the McKay Johns web scraping course a while ago and reimplemented most of the logic he taught there, but still, if you want to grab other variables you need to go over the process again by yourself.



After that, it gets a little less complicated since the mplsoccer and matplotlib documentation help a lot to understand how to build these plots. As a tip: if you combine your doubts, your context, along with the documentation, and you pack all that into a prompt for a LLM like GPT-5 or Claude, you'll get extra valuable help like I did. For this project, I've been iterating with Claude Sonnet 4 to help me generate custom designs for the dashboard and to help me debug Fotmob web scraping errors (the web that was harder to scrape in my opinion).

CJ: In your opinion, how will the role of a data scientist change within football over the next few years, and will AI change how people

approach football analytics looking specifically at scouting?

R: Like in most technology roles, the data scientist role is changing right now as we speak at a really fast pace in the industry, both in soft skills and technical skills. You're no longer expected to only train a model

inside a notebook and pass it to the ML engineers or software development team to deploy it, you're expected to go a little further owning the end-to-end process by being a "full-stack" data scientist depending on the maturity of the organization you're working for.

If inside your organization there aren't defined roles as data engineers, ML Engineers or developers, you might be expected to wear many hats and touch on several areas:

extract structured and unstructured data from different sources, store it, clean it, manipulate it, and retrieve it whenever possible (data engineering), doing EDAs, training models, making experiments, hypotheses test, build web apps (data scientist), and deployment at a working level (MLE, SW). I would also add to the

stack: having Generative AI literacy, since these technologies and solutions are helping developers solve problems that might have taken you days, but in minutes or a few hours. Besides that, you also need to improve or develop soft skills like communicating all your findings and results in simple terms, estimating development times, especially when you're dealing with non-technical stakeholders and clients that need the easily digested information.

In my humble opinion, this trend will go up since AI will help teams upskill and augment themselves (if they want to). It can be used to go one step further, helping you worry about the broader picture like: designing pipelines, end2end solutions, apps, strategies, and automated workflows, letting AI do the heavy lifting and scaffolding, always with you behind the wheel (this is important).

Regarding scouting: yes, it will empower curious individuals (analysts, coaches, scouts, developers, etc) be it freelancers or within clubs, agencies, and scouting departments to build better and more profitable solutions for them. These opportunities will depend on the knowledge and budget of your organization. If you have a high budget you'll have access to premium data and tools from top providers, and probably

you'll have the chance to have tailored AI solutions for your specific use case, that will give you an edge by helping you cover more leagues and players.

On the other hand, if you're on a tight budget you can also build your own custom solutions using open-source models and platforms to specialize in niche leagues and find great market opportunities for example, using it in other statements of clubs like youth academy development marketing departments, or any other use case you can come up with.

The important thing will be developing an AI-first brain and combining that with your domain knowledge to solve problems you care about.

CJ: Who are some others in the community that act as inspiration or are a great place to start when wanting to enter the industry in terms of upskilling in soccer analytics?

R: I mention some of them in the Sub stack post of the 5 Actionable Actionable Tips and Resources but it is worth mentioning them again and some others, since they've helped me along the process:

1. **[McKay Johns](#)** both YouTube Channels and football analytics affordable **[courses](#)**.
2. **[Ben Griffis](#)**: Advanced visualizations, he shares free apps for scouting and has a Google sheet to watch free football matches from less popular but interesting leagues for scouting.
3. **[Liam Henshaw](#)**: he created a community to help you get in the football scouting industry.
4. Jesús Lagos (**[@Vdot Spain](#)** on X/Twitter): he's a scouting consultant with a long-term solid experience working with top clubs and agencies. He shares great content on how to use data for scouting, going beyond just aggregated and event data, but creating your own metrics to better understand football actions and scout players with more context.

LOGGING OFF,  
CJ

## JOURNALISM TO DATA STORYTELLING – TANYA LOMSKAYA

Hi all,

I am joined by Tanya Lomskaya this week. We had the opportunity to talk through her wonderfully designed portfolio of work and a few tips when it comes to utilising a range of toolkits to create work.

Find our conversation below,

CJ: Tanya, thank you for giving up time to share more on your portfolio of work. What was your entry point into the world of data? Did your background in Economics and finance push you this way?

Tanya Lomskaya · 1st  
BI Consultant | Data Visualization Developer | Tableau Public Featured Author & 6x #VizOfTheDay |  
medium.com/@lomskaya  
Dubai, United Arab Emirates · [Contact info](#)

T: It is difficult for me to pinpoint a specific entry point. I received a degree in sociology but didn't pursue a career in it. I dreamed of becoming a journalist, and so after graduating from university in Minsk, I went to Moscow to make my dream come true. I became a journalist and spent eight years in large business publications, writing first about stock and insurance markets, then about macroeconomics, central bank and budget policy.

Business journalists who write for print can be divided into two types. The first are focused on establishing and maintaining relationships with sources—officials and top managers of companies—who leak information to them from time to time. The second work more with open data: they track business indicators and government statistics and interact with analysts and scientists. I belonged to the second type. If you want some parallel with the work of an analyst, then my first job was to identify outliers—events and indicators that stand out from the general trend—and try to understand why they happened, whether there had been some significant social or business event behind them. Such a statistical anomaly could turn out to be an unnoticed consequence of government cover-ups—in one case out of a hundred. In the remaining

99 cases, there is nothing worth nationwide reporting behind such outliers, and you spend tens of hours just to reapprove this simple truth.

I obviously needed tools to speed up the processing of information so that we could check the newsworthiness of statistical events more

quickly. I started learning Python and writing scripts to process government statistics, tax returns, and the like. It was quite fun, because some economists and researchers I was collaborating with also started coding at the time. It expanded their ability to do research so much, and there were some just phenomenal ideas there. I learned a lot from the people I was surrounded by at the time; they changed my view of working with data and generally my view of the world in many ways.

CJ: You have a wonderful eye for design, that goes beyond just Tableau visuals and the technical proficiency needed to create such elegant pieces. For your more recent “**Our Worlds Aging**” visual, how do you go about planning a map piece with data “layers” of the area charts and circles.

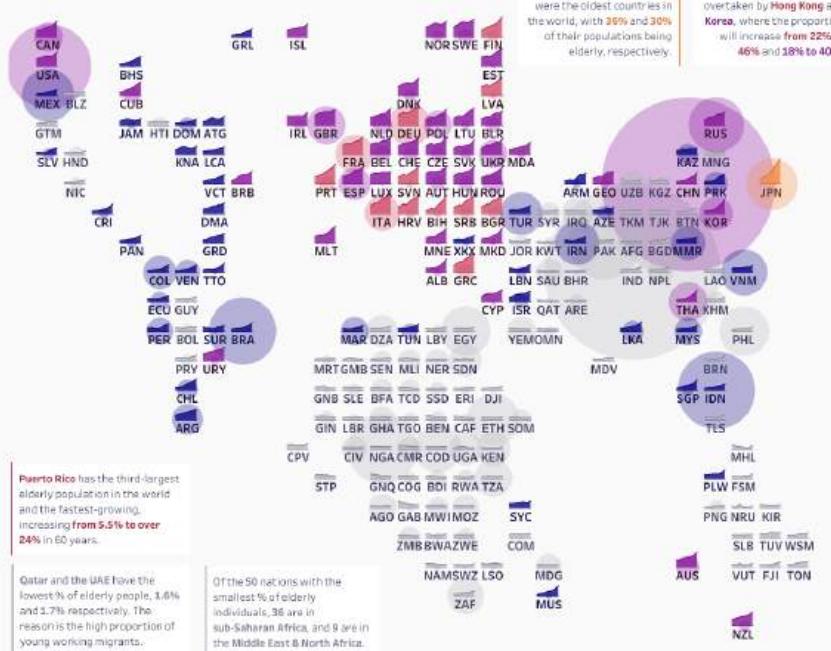
## OUR WORLD'S AGING SOCIETIES

We live in an aging world: since the mid-20th century, the global share of the elderly population (65+ y.o.) has doubled from 5 to 10 percent and is projected to rise to 17 percent by 2050. A population is generally considered aging when the proportion of older people exceeds 7%; today, in some countries, it has already approached 30%. Explore how societies around the world have aged over the past 60 years, and what has caused them to age.

### HOW TO READ THE CHART

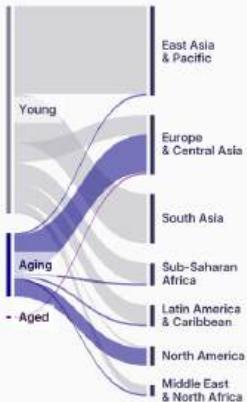


Hover over a country to see its age pyramid change.

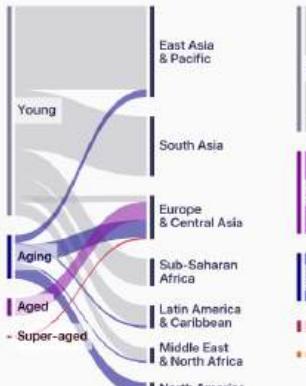


## AGING REGIONS

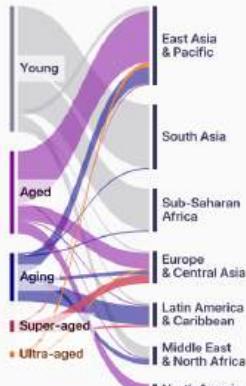
1963



1993



2023



Data: World Bank, UN

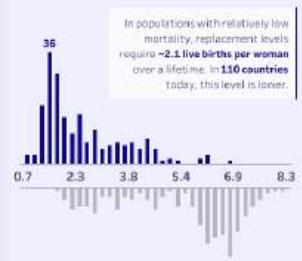
Sankey Generator: Ladataviz.com

Design: Tanya Lomskaya

## COUNTRIES' REDISTRIBUTION

### FERTILITY RATE 1962 • 2022

The global fertility rate has fallen from 5.0 births per woman in 1962 to 2.3 in 2022, resulting in a decline in the proportion of young people in the population.



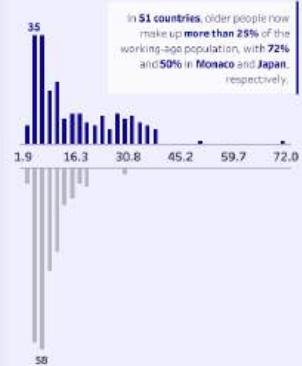
### LIFE EXPECTANCY 1962 • 2022

Life expectancy at birth increased by 17 years, from 55.2 to 71.9 years, leading to an increase in the number of elderly people.



### OLD-AGE DEPENDENCY RATIO 1963 • 2023

The modern working generation needs to support a larger number of elderly people: if in 1963 there were less than 9 people aged 65+ per 100 people of working age, in 2023 there were more than 15.



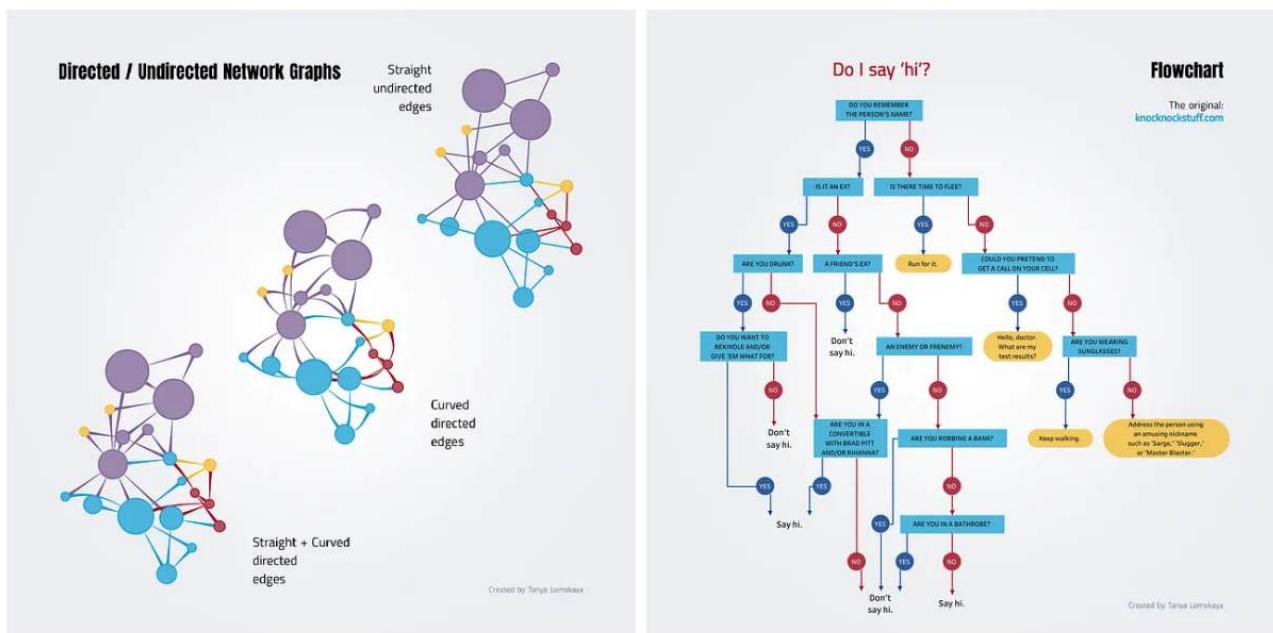
T: I borrowed the two-layer map model from Klaus Schulte's IronViz 2019 entry.

I was hooked by the ability to display the indicator as both a current static value (the area of the underlying circle) and a historical trend (the height of the overlaying bars), and all this with a geographic breakdown.

I decided to apply his approach to visualizing the aging of the world's population, with a slight twist. While the polygons show the dynamics of elderly population share since 1963, the resulting percentage is expressed not by the size but by the colour of the underlying circle, while its size represents the total population in the last year of observation. This allows us to see, for example, not just the significant current share of elderly people in China, but also how much China contributes to the global picture of aging, given its huge population.

I also added annotations to bring the demographic information to life and set context for the whole narrative.

CJ: From a technical aspect, I love that you're willing to share **more complex charts** in Tableau. Going beyond your medium blog on graphviz, are there any other charts that you'd like to replicate in the future? How do you go about planning non-standard chart types design?

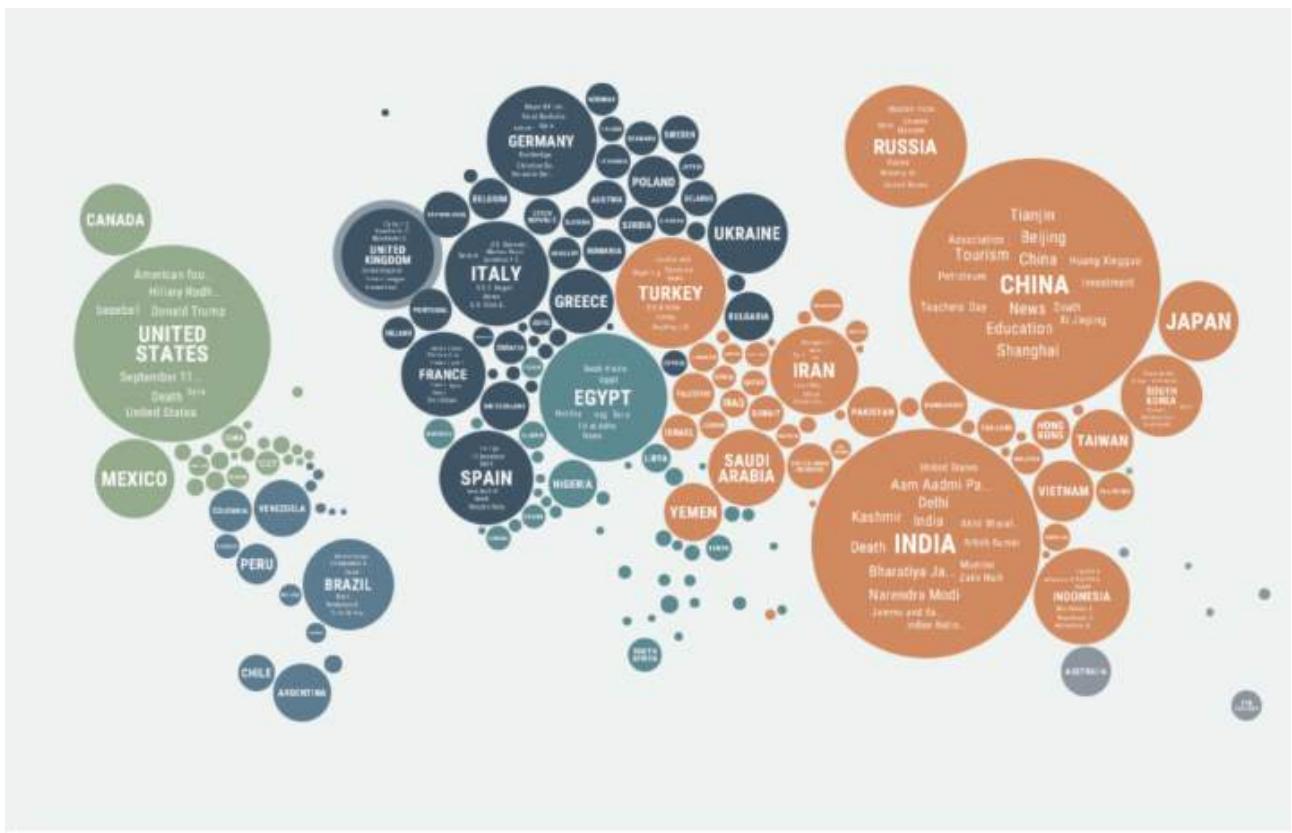


T: I think that the need for a non-standard chart design mostly arises when you want to depict some complex phenomenon as a whole, without breaking it down into separate indicators. So initially, the audience sees just one graph, which, upon closer inspection, breaks down into many details.

Take, for example, the **timeline** in the form of a curved area bump chart like this one from Robert Janezic. First, you see the timeline as a whole, along with a couple of genres that finished most popular; then

you change the lens and see the path of each genre individually. Then you focus on half-decade percentage distributions; finally, you dig into the story behind the annotations. I tried to create something similar, but so far nothing worthy has come out.

Also, I'd really like to find a way to create Dorling cartograms like this one from IIBA-2016, but in Tableau



This seems like a really challenging idea, because other than the one mentioned above, I've seen maybe one or two visualizations of this type that actually looked good and not just like unintegrated bubbles. So far, the best attempt in Tableau was by **Raisa Hannus**; you wrote about it last year.

Besides these, I'm fond of recursive **Voronoi generative arts**, like this fractal made with D3.js. I've written Python code to generate this kind of graph for Tableau but have not yet found a topic to use it on.

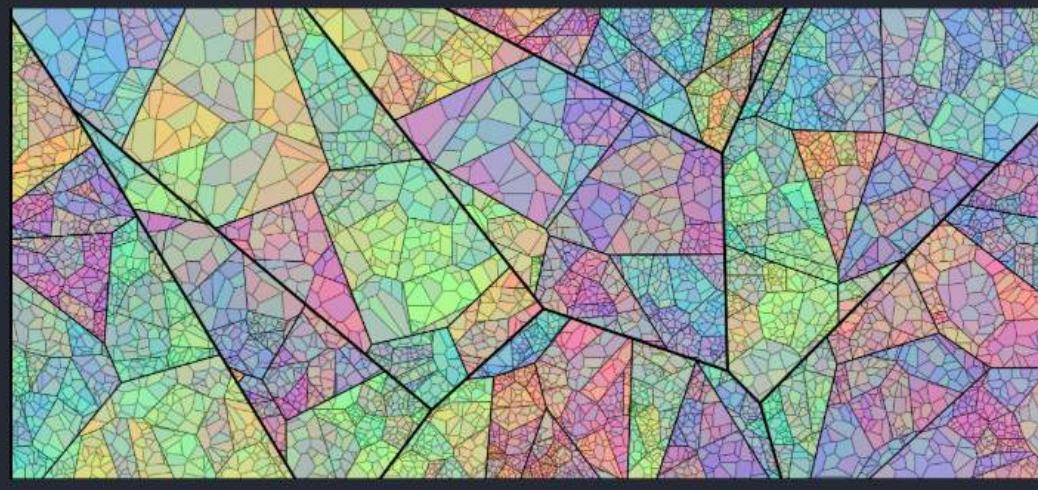
## Voronoi Fractal with D3.js

■ November 25, 2016 ⓘ less than 1 minute read

I've been very busy the last couple of months. Started a new project at Philips, started writing on D3.js for Packt, and besides that participating in two startups. So time to update this blog has been limited. While thinking about how to best present the voronoi feature of D3.js, I ran across an example where they used the result of a single voronoi partitioning as an input for another. So you'd get a nested set of voronois. This actually looked really nice, so I set out to do that myself, so I could use that example in the book as well.

### Results

The results look really nice (at least I think so):



CJ: Often for tableau visuals now in public, a whole range of tools get used to create them both for inspiration, data prep and design. Do you have a few "go-to's" you can share that might help others elevate their own visuals?



T: If you love creative visualizations, then your path in any case lies in the direction of D3.js or Python. For Pythonists, I can advise paying attention to the Shapely library for generating and transforming polygons, SciPy BPoly for generating curves, and PIL for image processing.

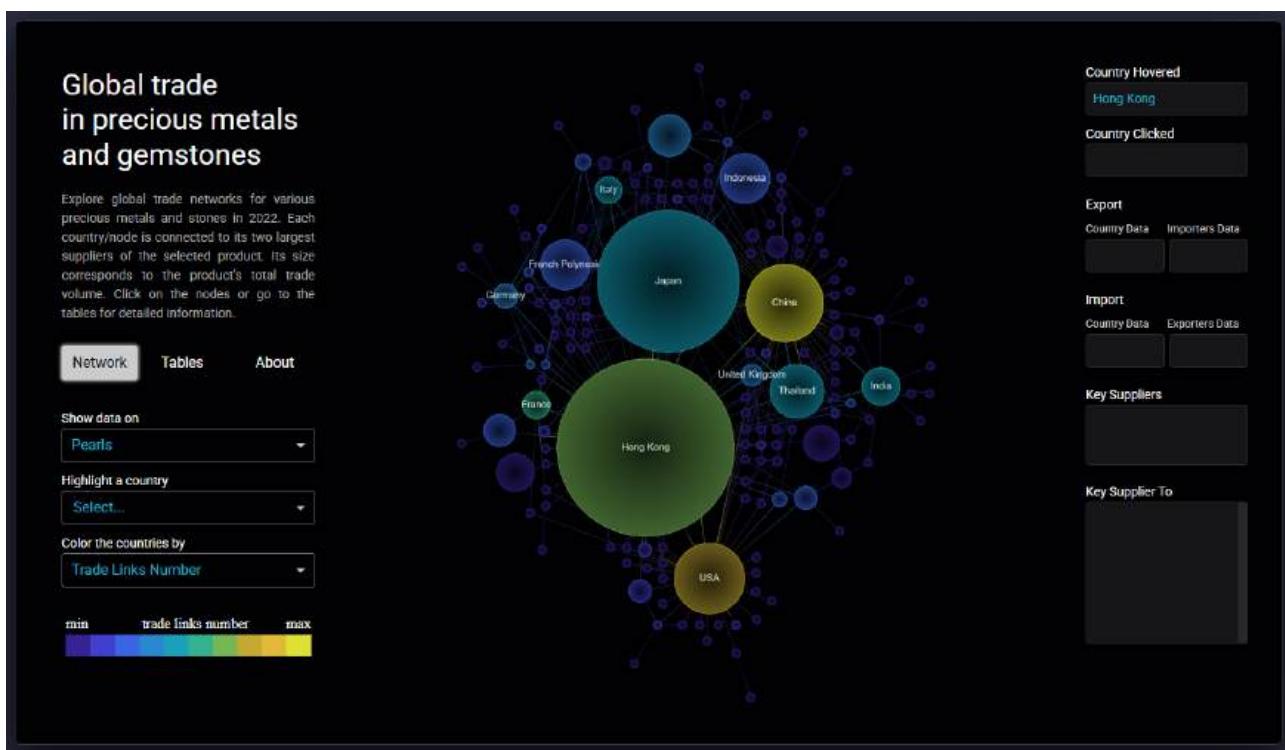
- > **Shapley**
- > **SciPy**
- > **Pillow**

If you're not yet a coder, my go-to is to become one.  
(Just kidding.)  
(Or not.)

I also always recommend studying Inkscape, an open-source Figma-like package, but more oriented towards visual design. The most basic thing you can do with it is to extract the SVG file with labels from [Ladataviz.com](#), change their colour, font, and size, and then upload the resulting image to Tableau as a background for your chart. You can do a variety of more advanced text tricks, generate visual effects for your background and navigation elements, and create the most intricate shapes—there are plenty of tutorials on this topic on YouTube.

CJ: You've done some work on dashboard design using other tools and technology, like your **precious metals** piece. You reference some places to help you get started with Dribble, Dash, Plotly and

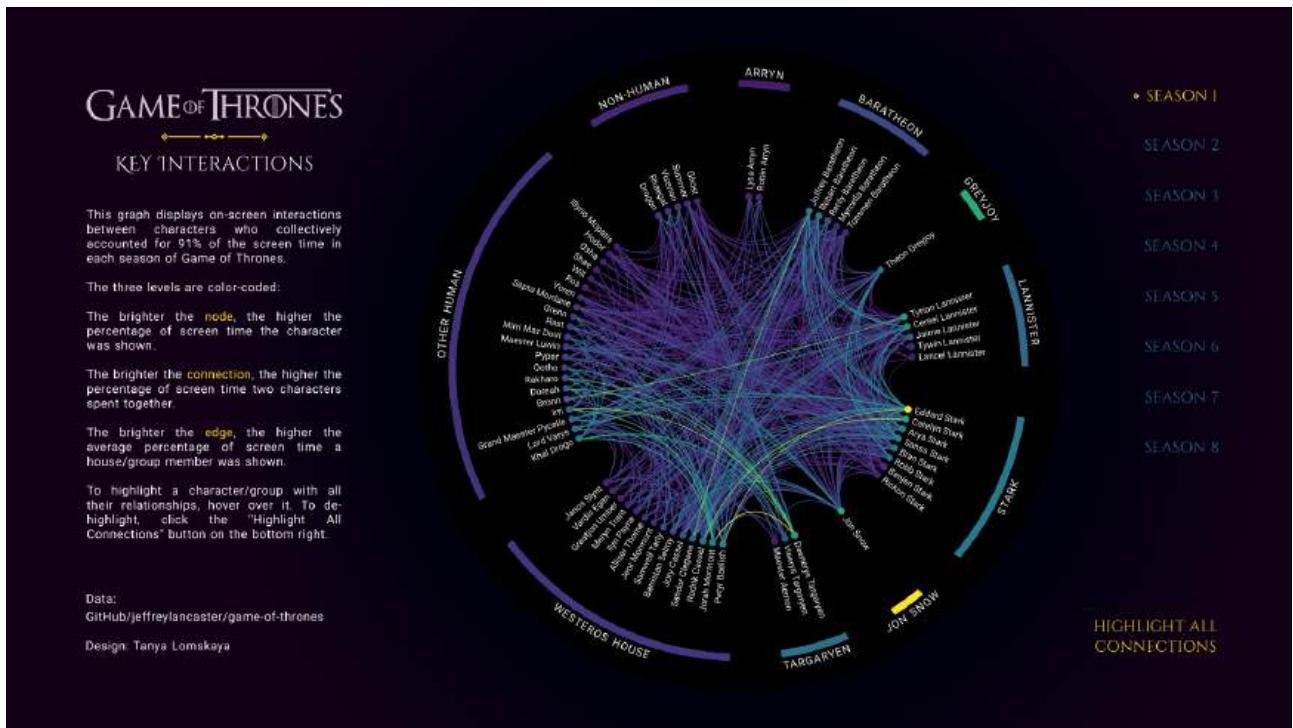
stackoverflow. With advances in AI and prompts, what build aspects will become easier?



T: This will change the whole process of creating apps. I tried creating an app in **Plotly Studio**, which is currently in beta testing; the experience reminded me of working with Midjourney. You don't need to know coding at all.

It feels like dashboard/app creation is changing its nature. Instead of developing from scratch, we will improve in interacting with AI, while our technical skills will drift towards revising and correcting AI creations.

CJ: Chord charts are not easy! Your visualization on GoT covers all 8 seasons with colour-coded indicators – what design challenges did you face in representing such an extensive timeline in a single visualization while maintaining clarity about how character interactions evolved over time?



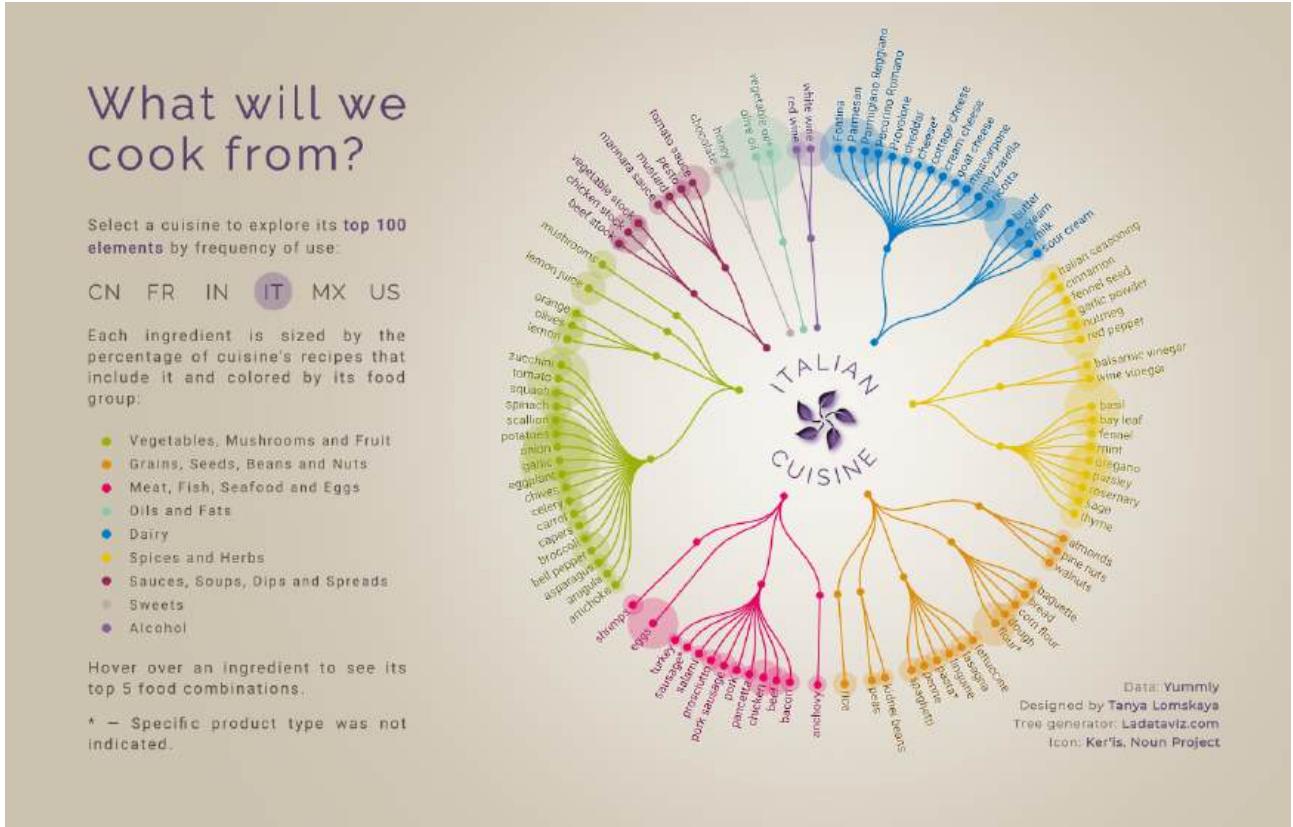
T: 90 percent of the entire creative process in this case happened outside of Tableau. I calculated the coordinates for all three layers of the chord—nodes, links, and edge polygons—using Python. (By the way, I learned to build all these elements in Python thanks to a cool tutorial on radial bar charts from the Flerlage twins: I simply replaced the Excel functions from their step-by-step manual with Python ones and began to calculate the coordinates for all my radial graphs.)

I created all the labels for the chord using Ladataviz.com and Inkscape.

Then I turned to Tableau to connect it all, attach the necessary backgrounds to the necessary graphs, and set up navigation between the elements. The final design is simple: in addition to the chord itself, there are toggle buttons for switching between seasons and instructions on how to use the graph.

The main challenge here was the amount of preliminary work: developing the code and generating multiple background images.

CJ: Your circular or chord visuals “Who’s In The Spotlight?”, “My Movie Color Wheel” & “What Will We Cook From?” are all the centre focus of the visual. Is this a conscious decision? How do you weigh up what other charts or text to align on the page so as to not detract from the main visual.

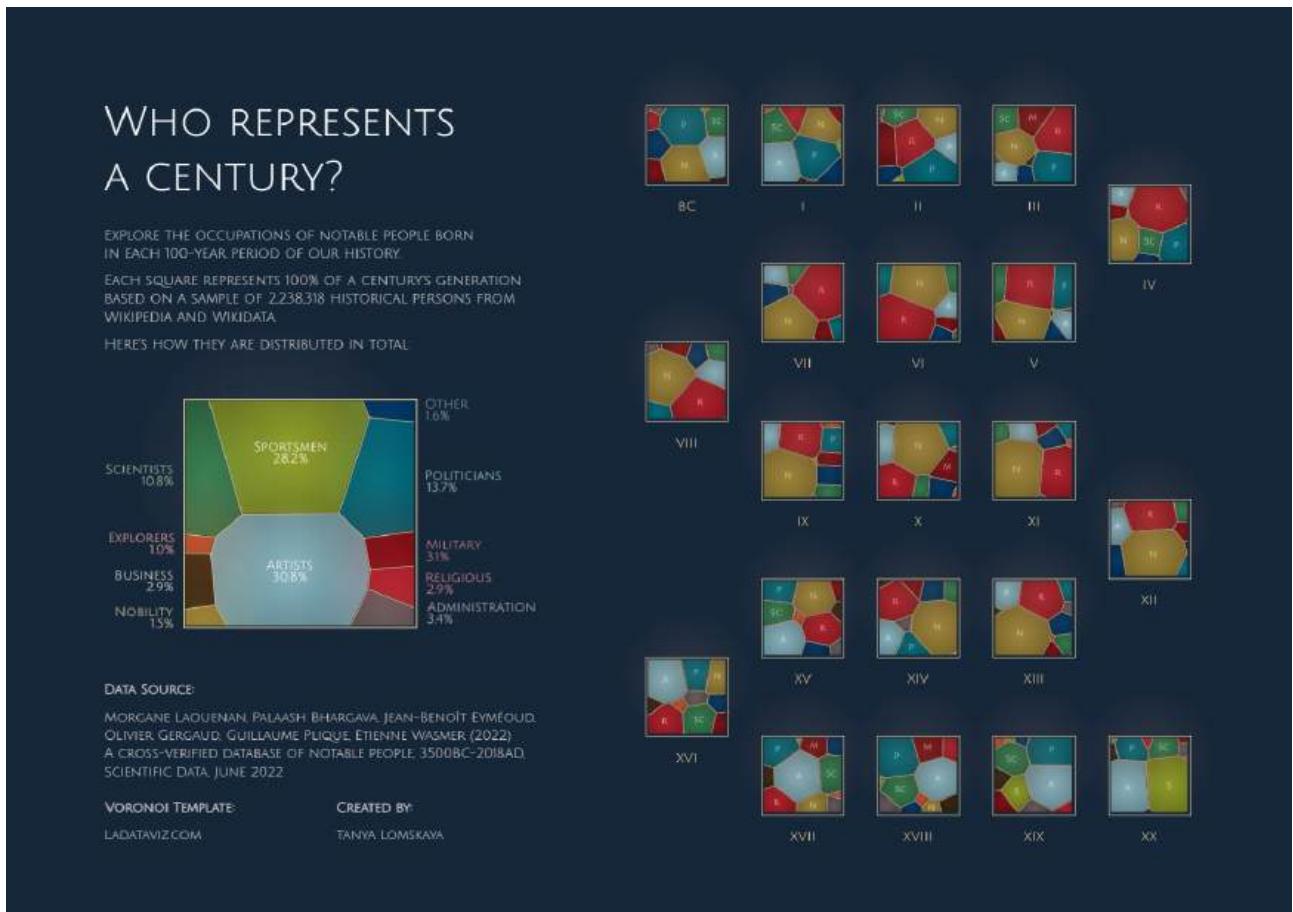


T: Everything that goes beyond the circle is perceived as the periphery—hence this design. Almost all important information is contained within the circle, and everything outside is either legends, navigation elements, or auxiliary charts. This approach seems very common; I have seen many authors utilizing it. Whenever I create a radial, I place it exactly like this.

CJ: What has been your favourite visual to make and why?

T: If you mean which of my visualizations left the most positive emotions while developing, it's "**Who Represents a Century**." It was a purely creative process from start to finish, more like a game, based on a very interesting dataset. All the way I was focused on the visual part and creating a certain perception that I needed. And the cherry on the cake is how warmly the result was welcomed by the community.

CJ: What guided your decision to use Voronoi diagrams to represent occupational distributions rather than more conventional charts, and how does this visual approach enhance the storytelling about historical representation?



T: It was done more on a whim, but here's how I explain the idea. Sciences Po cleaned up and systematized Wikipedia data on people of different professions who lived in different eras, up to the present day.

But Wikipedia is not an objective source of knowledge about the inhabitants of a particular time. It contains information only about those people who, for some reason, remained in the memory of mankind.

In other words, this data is about our perception or fantasy about the past, not about the objective truth. Thus, the idea was to simultaneously show you each century's proportions and somehow make you doubt them.

Voronoi came in perfect for this. Apart from the stated effect, this graph resembles a mosaic or stained glass, which is strongly associated with history, or abstract art, which completely appeals to our subjective perception. To enhance this impression, I gave the graphs the appearance of an art gallery.

CJ: In the wider data visualisation community, are there any portfolios of work that have really inspired your journey, or made you think “wow I'd like to make something like that!”?

T: There are actually a lot of people about whose work or entire portfolios I think, "Wow." Let me list only those who have significantly influenced the creation of the visualizations already published in my portfolio. These are the Flerlage twins, Andy Kriebel, Chimdi Nwosu, Priya Padham, Pradeep Kumar G, Kevin Wee, Dennis Kao, Wendy Shijia, Klaus Schulte, Will Sutton, and Damola Ladipo.

CJ: Are you working on any exciting personal projects?

T: I have been wanting to start making maps for a long time now, as I am incredibly inspired by the work of Dennis Kao, Nir Smilga, Naresh Suglani, and George Koursaros. Next, I want to continue the theme of representation of people in history, but this time focusing exclusively on women. I am curious what the data will show. And at some point in the future, I would like to do a project dedicated to my home country, Belarus.

Have a great week ahead!

LOGGING OFF,

CJ

#### GARMIN ANALYTICS – RUTH POZUELO MARTINEZ

Hi all,

I am joined by Ruth this week, who has been giving back to the community some amazing fitness garmin resources.

She was kind enough to have a chat with me about it in more detail.

Find our conversation below,

CJ: Before we jump in. Am I right in assuming you got into data through boutique BI consulting?

R: I think I have always been into data, but it was when I started Curbal (<https://curbal.com>) that I could do that for a living.

Welcome to Curbal!

Are you looking for resources to learn Power BI? Then you come to the right place.

Check out the video for more information about all Curbal resources available to you.

BECOME A MEMBER    SUBSCRIBE TO YOUTUBE

You can find us also here: [YouTube](#) [LinkedIn](#) [Medium](#)

CJ: What is it that you find so exciting about Matplotlib. I see you managed to recreate quite the famous “1 dataset, 100 visualisations”. How does the flexibility suit some of those graph types? Were there any stand out visuals that were particularly challenging? Thank you for sharing the code with us!

R: You can do anything you want with matplotlib. You can control every single pixel on the visualization and for me it doesn't get any more powerful than that. I can not blame it on the tool, if my visualization is not good it is because of my data/design skills, there is no hiding... haha!

You are here: [Home](#) / 1 dataset, 100 matplotlib visualizations

## 1 dataset, 100 matplotlib visualizations

At the beginning of the year I challenged myself to create all 100 visualizations from the [1 dataset, 100 visualizations project](#) and I am sharing with you all the code for the visualizations.

Note: Data via Project, or copyright Pelego and available under a [Creative Commons Attribution - Non Commercial - No Derivatives 4.0 International](#) license. I asked Pelego and they used a Design tool to create them, I did not ask for help.

**How do I get the code?**

Click on the image to go to the tutorial.

**Collaborate!**

I still have a few to create (around 5) and some of the plots can be further automated. Also, I need to improve the code and some plots are missing details like legends or data labels.

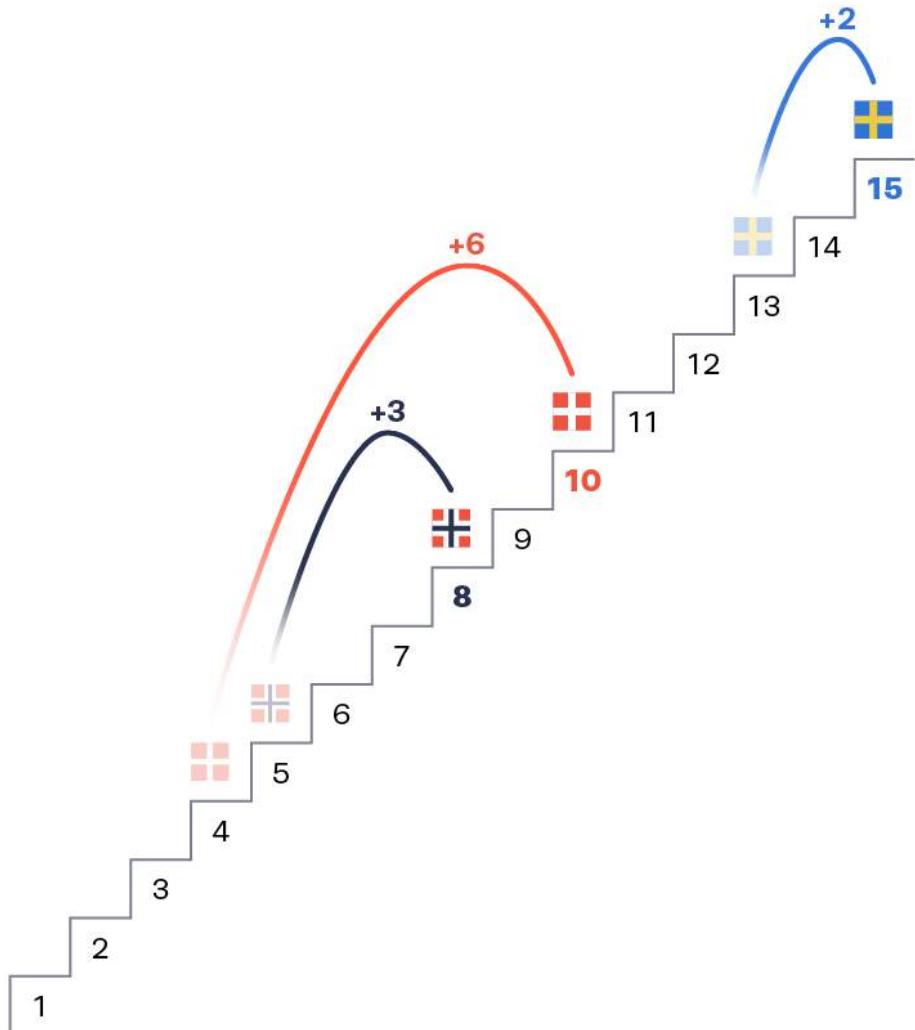
Feel free to tell me in the comments of each tutorial if you managed to improve the code in any way and I will update it with your changes.

**The matplotlib gallery**

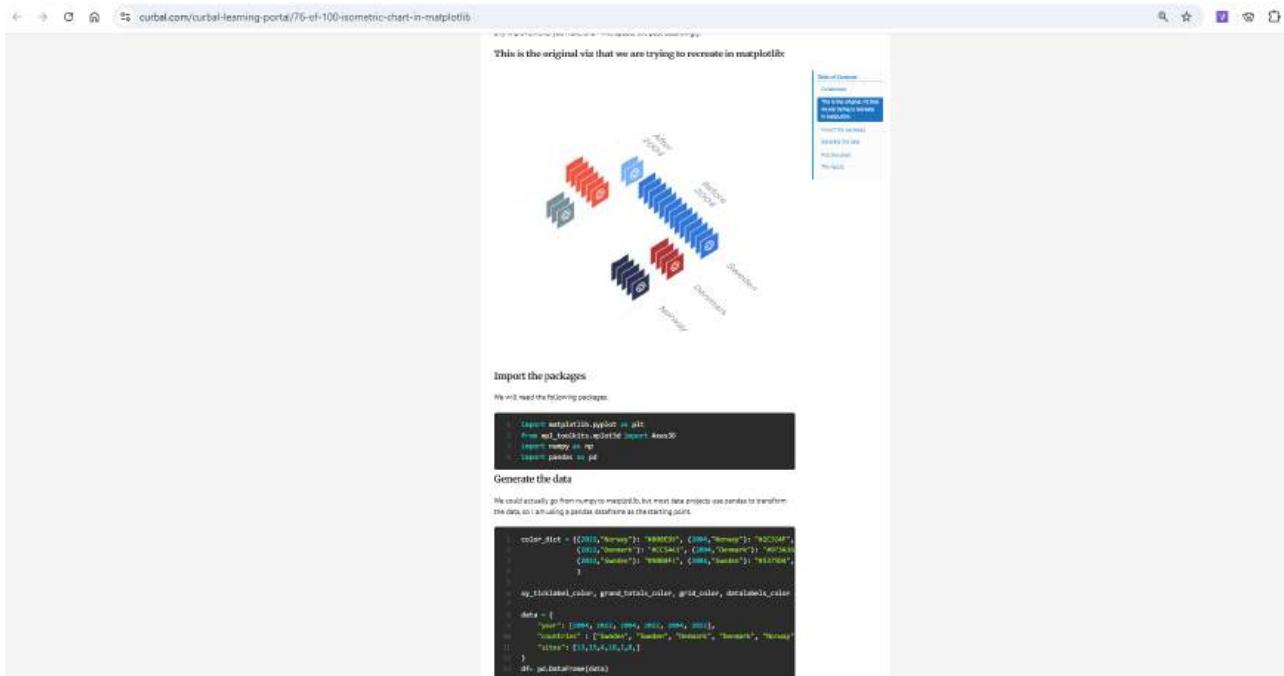
[Area](#) [Bar](#) [Dot](#) [Dumbbell](#) [Histogram](#) [Dotplot](#) [Piechart](#) [Scatter](#)  
[Slope](#) [Stair](#) [Stack](#) [Tornado](#) [Volvo](#)

The challenges I still have left for the 1 dataset, 100 matplotlib visualizations (<https://curbal.com/1-dataset-100-matplotlib-visualizations>) are to curve text and to pack circles. It is possible, but it is not included in matplotlib by default making it harder for my current skills to do.

CJ: Here are a few of my favourites from Ruth's work in the collection.  
(40 & 76)



With number 40 I love the step ladder and gradient jump / leap. It's a nice play on progression and a more basic arc chart. The movement up and right simultaneously really adds to the leap effect.

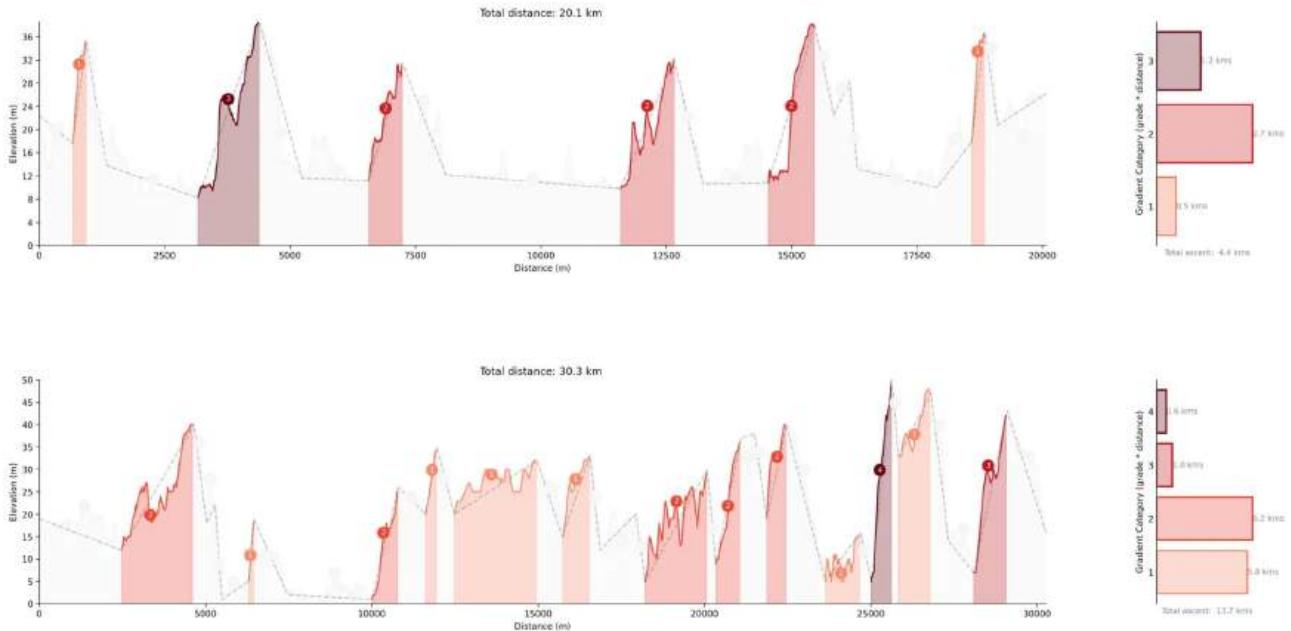


My other favourite is this isometric design. I like the uniformity to it, almost like a filing cabinet?

CJ: We have seen a large shift to people wanting to use “tools”. Your elevation profile plot allows just for this, It’s great you make it available for everyone. What are some of the complexities behind making something like this?

R: I am hosting the tool at pythonanywhere (<https://eu.pythonanywhere.com/>) and they make this stuff as easy as it gets. I am in their hacker plan, which means not a lot of compute or space so I can't keep those files in the app, but even if I could, for privacy reasons I wouldn't do it.

CJ: Recently you've shared a little around new projects on elevation profiles, through garmin connect data. For those looking to get started with Garmin connect data, where would you recommend starting?



R: This one lets you compare the elevation profiles of two GPS or GPX tracks — really useful if you're training for a specific event and want to see how your training route stacks up against the actual race course.

In short it,

Simplifies the tracks using the Ramer–Douglas–Peucker algorithm to calculate total ascent/descent and the difficulty of climbs and descents. Difficulty is measured by multiplying the slope grade by the distance — a value of 1000 marks a Category 1 climb, 2000 is Category 2, and so on.

Choose slope type (ascent or descent) using the dropdown.

Segments setting controls how much the track is simplified. A lower number = more detail. Start with 10 and tweak as needed.

Processing time is about 30 seconds for an average-sized GPX file.

### Check it out [here](#).

The other project I created is if you want to visualize your activities on a

3D elevation map, you can do it now [here](#).

CJ: Note as its a gpx file this will also work with exporting your strava data!



Made with ❤ by  curbal

Garmin does not have a public API, so the “easiest” way to get the data is use a python package called garminconnect

(<https://github.com/cyberjunky/python-garminconnect>) . I do have a developer account with Garmin, so hopefully by the end of the year people will be able to connect their own data to get my visualizations.

Fingers crossed!

CJ: When we caught up, you walked me through some of your personal garmin metrics for aerobic and anaerobic training, and told me how you automated the refresh process. Could you give a little context as to the steps involved in something like this through the data retrieval, capture, transformation and then visually displaying it on your site? (**Pillars of Endurance Running**) What tech stack would someone need to know?

# Running Dynamics

The science of running. Learning how to improve my running technique.

## The science of running

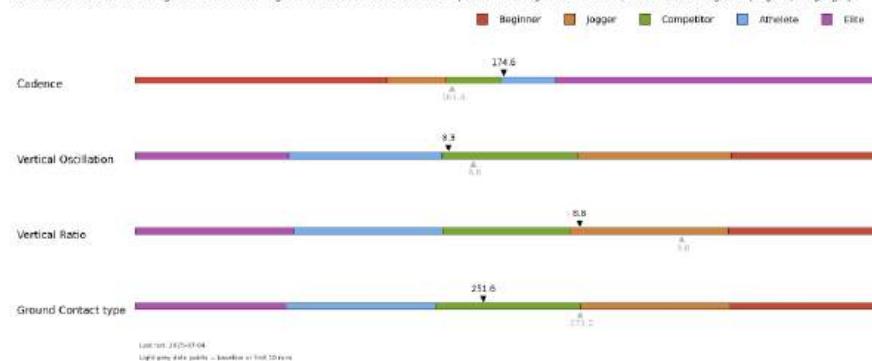
Garmin Epix watch provides 4 metrics: (click on each link to learn more)

- » Cadence
- » Vertical Oscillation
- » Vertical Ratio
- » Ground Contact Time

Below you can see my current status for all four.

For more details on each category, click on the links above.

Note: The data shows the average of the last 10 running activities that are not Anaerobic or Sprints in black against the first 10 (when I started the garmin program) in light grey.



R: For public use I recommend using the garminconnect python package (<https://github.com/cyberjunky/python-garminconnect>) and then use my own server to schedule the scripts that send the data to a database. From there I have other scripts that grab the data and create the visualizations. Once they are ready, I use the pythonanywhere api to push the images to the app. I was doing everything on pythonanywhere but it takes too long to produce the visuals, so this was a better approach.

So the tech stack would be: python (mainly, flask or django and matplotlib) and sql

CJ: It's exciting to see you are training for Lidingöloppet 30km using the Garmin coach. How has revisualising all this data helped support your training? Has it had significant improvements on your performance?

R: Oh yes! I have never been faster.

I realized a few things with Garmin coach: there is an efficient way to run and you should try to learn it. Garmin data will let you know if you get better at it (<https://www.garmin.com/en-US/garmin-technology/running-science/running-dynamics/>) Knowing what your aerobic and anaerobic threshold is, helps planning the races and training. The data also lets you

know how much water you lost, the calories for the run, things like that that helps you plan similar runs and food intake better.

The screenshot shows two tables of running data from the curbal.world/anaerobic\_training website. The first table is titled "Anaerobic" and the second is titled "Sprints". Both tables include columns for Date, Score, Duration, Elevation Gain, Avg Max Pace, Max Cadence, Aerobic Benefit, Training Effect, and Execution score per lap (labeled 1 through 10). The data is color-coded by score (e.g., green for 91%, red for 59%).

Date	Score	Duration	Elevation Gain	Avg Max Pace	Max Cadence	Aerobic Benefit <sup>10</sup>	Training Effect	1	2	3	4	5	6	7	8	9	10
2023-03-23	91 %	60 secs	1m	3:53 (3:13)	217 bpm	Improving	Anaerobic	100	100	100	100	100	100	100	100	100	100
2023-05-01	92 %	40 secs	0m	3:47 (3:10)	219 bpm	Maintaining	Base	100	100	100	100	100	100	100	100	100	100
2023-04-22	93 %	40 secs	1m	3:49 (3:22)	221 bpm	Maintaining	Base	100	100	100	100	100	100	100	100	100	100
2023-05-15	78 %	40 secs	0m	3:50 (3:17)	222 bpm	Maintaining	Tempo	44	100	79	30	49	77				
2023-03-15	74 %	40 secs	0m	3:50 (3:09)	222 bpm	Maintaining	Tempo	49	57	49	52	44	54	77	87		
2023-03-08	71 %	40 secs	0m	3:50 (3:18)	222 bpm	Maintaining	Base	69	79	69	67	12	37	89	59	82	0
2023-02-22	59 %	60 secs	15m	4:31 (4:08)	219 bpm	No effect	Tempo	50	0	0	0	96	25	11			

Date	Score	Duration	Elevation Gain	Avg Max Pace	Max Cadence	Aerobic Benefit <sup>10</sup>	Training Effect	1	2	3	4	5	6	7	8	9	10
2023-07-02	06 %	15 secs	0m	3:15 (2:57)	229 bpm	Maintaining	Base	100	100	100	100	0	100	100	100	100	100
2023-06-19	92 %	10 secs	0m	3:20 (3:49)	229 bpm	Maintaining	Base	100	100	100	100	100	100	100	100	100	100
2023-06-06	75 %	15 secs	0m	3:27 (3:08)	217 bpm	Maintaining	Base	100	100	100	100	100	100				
2023-05-29	99 %	10 secs	0m	3:05 (2:53)	231 bpm	Maintaining	Base	100	100	100	100	100	100	100	100	100	100
2023-05-24	98 %	10 secs	0m	3:01 (2:53)	231 bpm	Maintaining	Speed	100	100	100	100	100	100	100	100	100	100
2023-05-23	96 %	15 secs	0m	3:17 (2:59)	223 bpm	Maintaining	Tempo	100	100	100	100	100	100	100	100	100	100
2023-05-22	94 %	10 secs	0m	3:09 (2:49)	233 bpm	Maintaining	Base	100	100	100	100	100	100	100	100	100	100
2023-03-22	97 %	10 secs	0m	3:15 (2:55)	240 bpm	Maintaining	Base	100	100	100	100	100	100	100	100	100	100
2023-02-12	78 %	15 secs	1m	3:49 (3:24)	229 bpm	Maintaining	Base	100	0	0	100	0	0	0	0	0	0

I could go forever, but to answer your second question, I just finished a project that clusters all runs made on the same/similar track so I can check if i am getting faster.

I have picked a few that are relevant for the lidingöloppet and it looks like I am indeed getting faster. That chart is at the end of the page for the anaerobic runs ([https://www.curbal.world/anaerobic\\_training](https://www.curbal.world/anaerobic_training)). It also updates automatically as I do more runs on those tracks, so I just need to focus on running faster!

CJ: What is next for you in the health and fitness data world?

R: I have soooo many projects I want to do:

1. I want to have a better follow up of the Garmin coach program
  2. Create my own “year in review” (this is one step towards that: [https://www.curbal.world/gpx\\_poster](https://www.curbal.world/gpx_poster))
  3. I want to be able to profile the race courses better so I can plan race day better. The elevation comparison is part of that work.
  4. I want to be able to analyze my running technique starting with my vertical oscillation. Python has great packages for that....
- The list goes on and on!

As a fun side note, remember that I told you my partner calls me Ivan Drago?

**Here is why.**

CJ Round-up: It was so much fun chatting with Ruth. What better way to spend time than nerding out over data, and also talking about a shared love for running and passing stories on some of our training plans (or lack of in my case!)

I hope the above conversation gets you excited to explore some of your own health metrics, be it through garmin or beyond – One thing I know is I'm excited to see how some of these new projects formulate in coming months for Ruth.

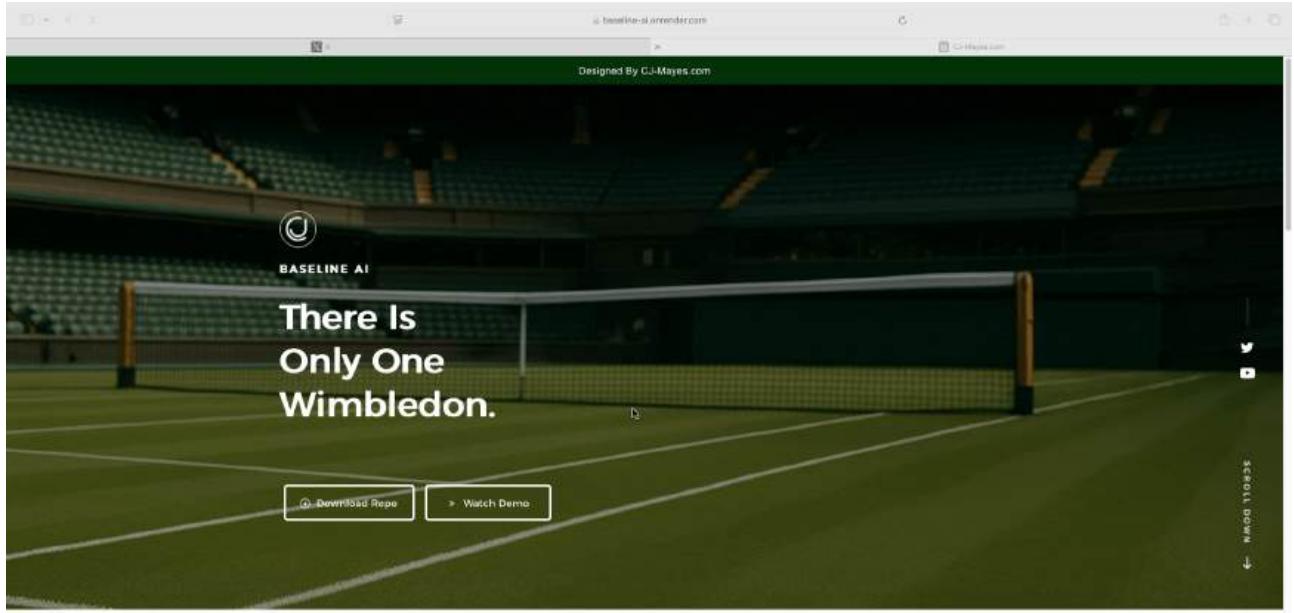
LOGGING OFF,

CJ

BASELINE AI – TABLEAU LANGCHAIN

Hi all,

AI is gold as I'm sure you're aware by now, and many of us probably have gut wrenching feeling that you need to grow with the times or be left behind. I'm particularly excited to see how tools are growing at pace to be able to not just bring efficiencies in other walks of life, but how it is reshaping the way we do analytics.

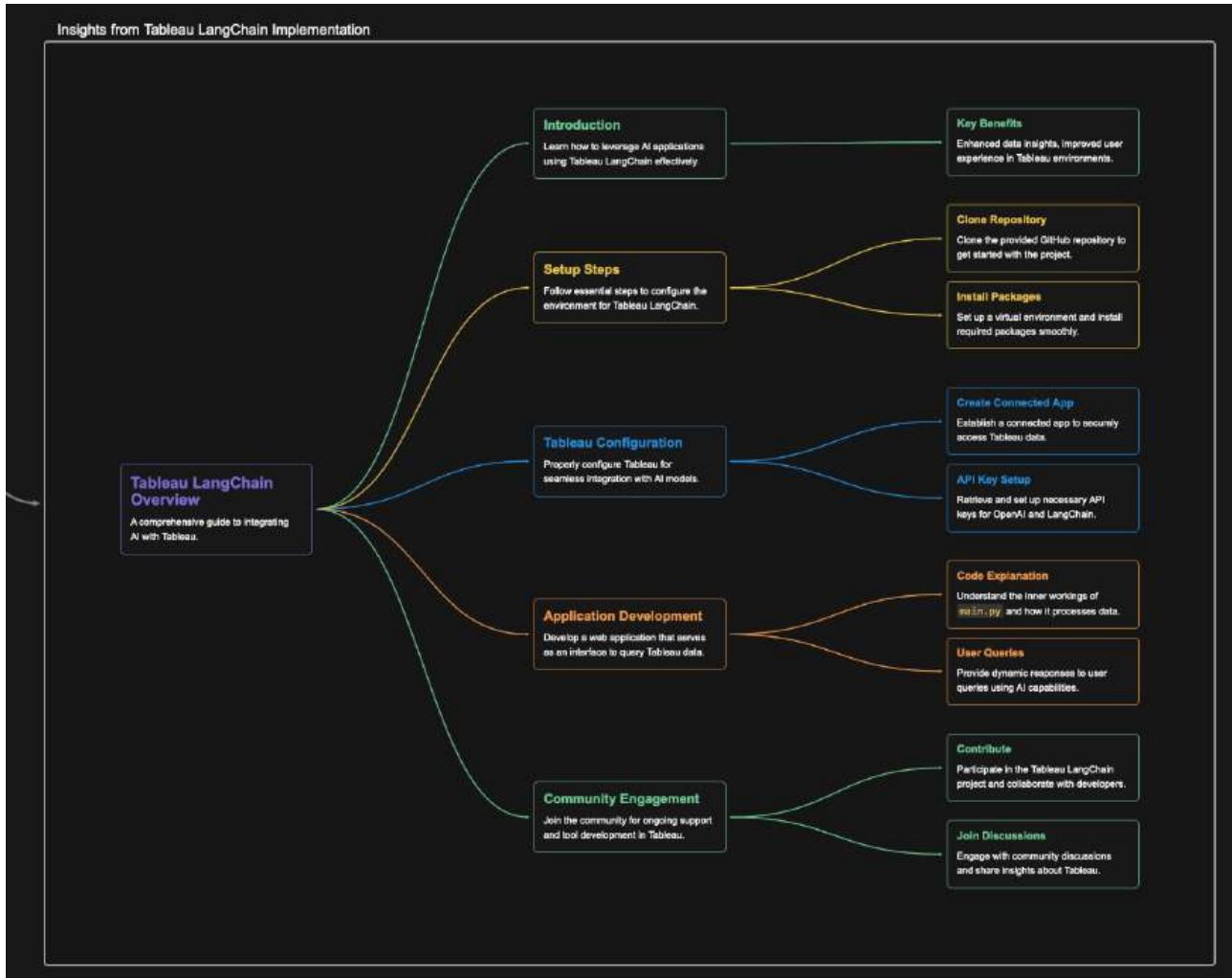


As part of my own journey, I want to expand a lot of my technical skills into the peripherals of data, that will include front end design, hosting, back end, AI, and everything in between. To do that I tend to set myself mini projects. One recently that came up for me was to play around with Tableau Langchain.

If you have no idea what that is, start here with Will Sutton. I got to see it live in action at AgentForce a few weeks back.

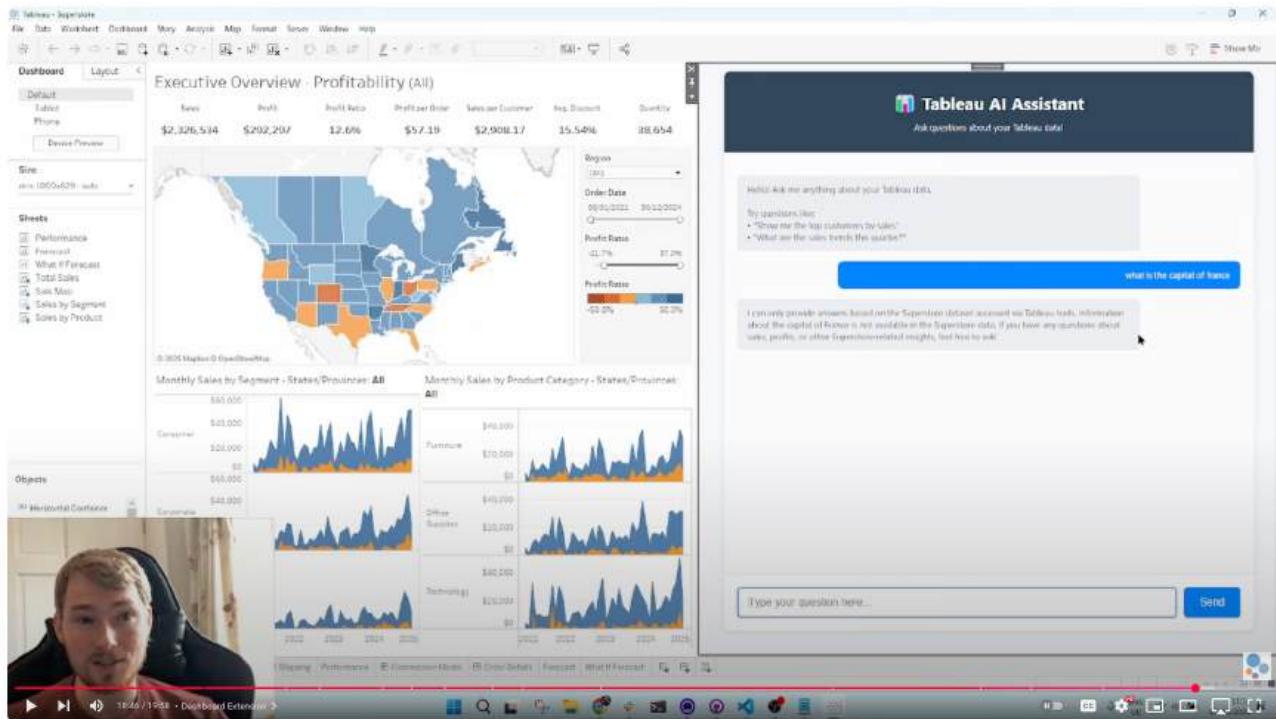
The video thumbnail features a portrait of Will Sutton, a Tableau Visionary, wearing a dark t-shirt with a 'dbt' logo. The title of the video is 'Add AI to Your Tableau Server (or Cloud) With Tableau LangChain'. The channel name 'The Information Lab' is visible at the bottom left. The video player shows a progress bar at 0:00 / 14:58 and various control buttons like play, pause, and volume.

This is everything you need to get started with the idea of querying your Tableau server dataset through prompts using an OpenAI model.  
Here is a short summary of the video



The video goes into great detail as to how to connect the Tableau side for the server, the open ai model, how langchain works and best way to utilise the repository.

I'm going to trust you also use this as a starting basis and then will talk through some of my next learnings.



So if you follow along you get a great chat bot that is an html page, that you can either bring into your dashboard or host locally given the repository.

The next steps for me therefore, in terms of learning by doing was.

- Replicate what Will has done using his repository
- Amend it to a sport of my choice, a different dataset and different agent style.
- Start adding in custom features that make it more my own project.

Let me walk through some of those adjustments.

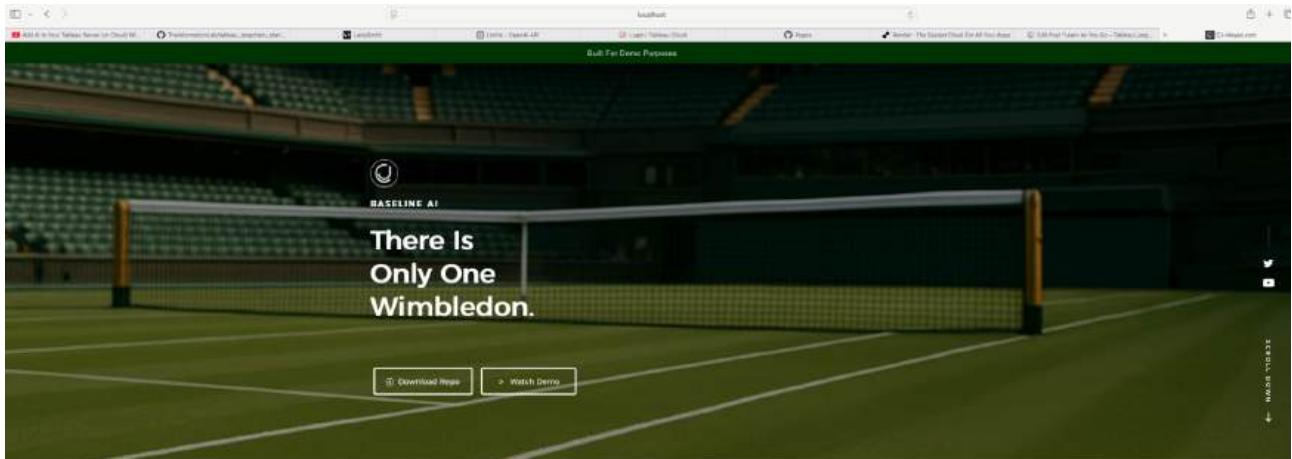
```

index.html
static
utils
env.py
env_update.py
LICENSE
main.py
Profile
README.md
requirements.txt
web_app.py

```

Firstly, I started by cloning the repository in cursor. I just love its ability to use the AI feature to query my whole repository instead of having to copy and paste code chunks out.

Next, I wanted to change the page to fit my new theme, which was going to be tennis data. This is mainly just changes in the HTML.



#### About Baseline AI

This is a simple AI chatbot that uses the OpenAI API to answer questions about Wimbledon 2012 data. It's created as a playground for me to learn about AI, Taboola, and front end development.

I got far too excited with this stage, playing around with different generated images, layouts and styles. I only picked up HTML recently after collaborating with Chelsea Football Club.

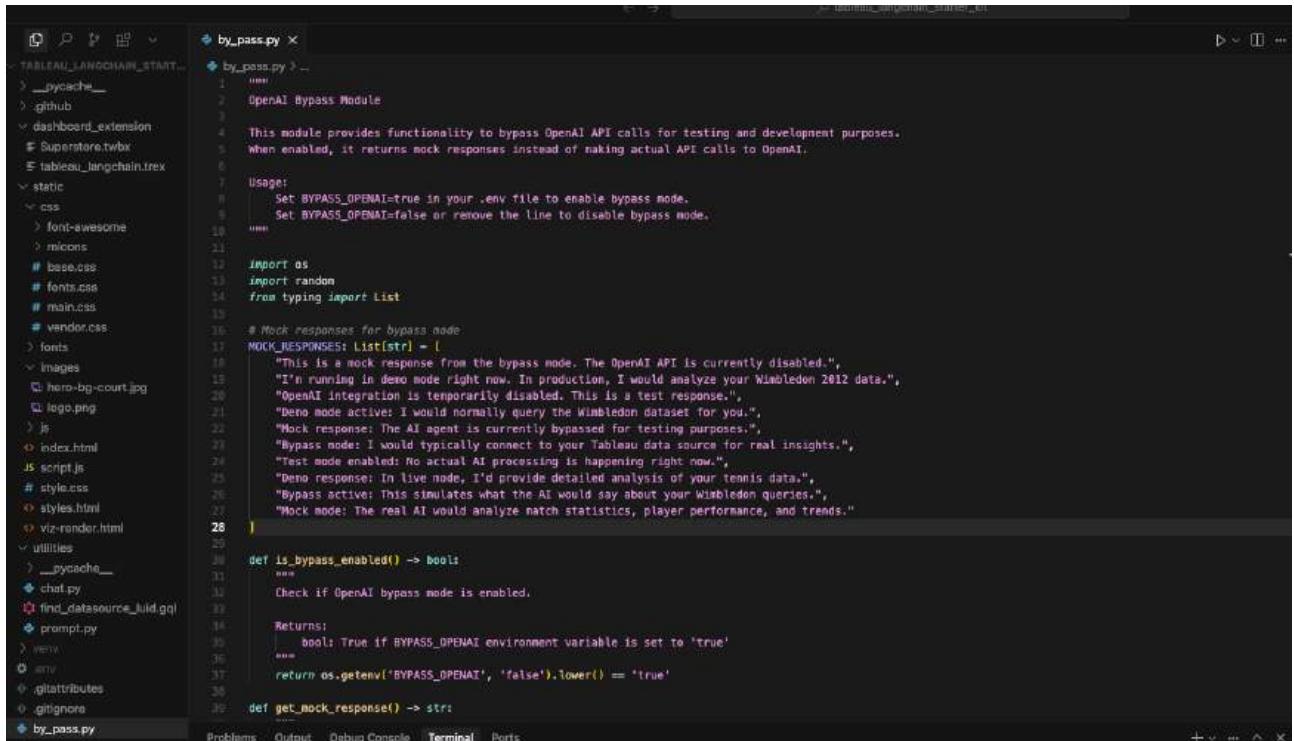
The next part of that is the “chat bot” itself.

A screenshot of the Baseline AI Chatbot interface. At the top, there is a "Data Download" section with a "Match Results" button and a "Point By Point Data" button. Below this is a "Baseline AI Chatbot" section with a note that the dataset was used for visualization. There is a text input field containing "Ask a specific question about Wimbledon 2012 match results...". To the right of the input field is a "Send" button. On the right side, there is a "Featured Prompts" section with four options: "Show me the match summary for the women's final.", "List of players who served more than 10 aces.", "Who had the highest first serve percentage?", and "How many points in a row did Federer win against Saïf�Erradi?".

I wanted the flexibility to be able to put in custom prompts, a nice guide so that users were able to easily identify common questions. It would also allow me to start to be able to really configure what those responses may look like. The other aspect was I wanted the flexibility to

download the data. Now really this should be actioned from the server, but for now I linked it to Jeff Sackmanns repo where I took the file from in the first place!

You may also notice that clicking these prompts gives you a bunch of nothing response at the moment. I created a file with a toggle in the environments variable to prevent the message hitting the model (because costly) – Instead it just loops through these generic responses for now.



```
by_pass.py
by_pass.py
OpenAI Bypass Module
This module provides functionality to bypass OpenAI API calls for testing and development purposes. When enabled, it returns mock responses instead of making actual API calls to OpenAI.

Usage:
    Set BYPASS_OPENAI=true in your .env file to enable bypass mode.
    Set BYPASS_OPENAI=false or remove the line to disable bypass mode.

import os
import random
from typing import List

# Mock responses for bypass mode
MOCK_RESPONSES: List[str] = [
    "This is a mock response from the bypass mode. The OpenAI API is currently disabled.",
    "I'm running in demo mode right now. In production, I would analyze your Wimbledon 2012 data.",
    "OpenAI integration is temporarily disabled. This is a test response.",
    "Demo mode active; I would normally query the Wimbledon dataset for you.",
    "Mock response: The AI agent is currently bypassed for testing purposes.",
    "Bypass mode: I would typically connect to your Tableau data source for real insights.",
    "Test mode enabled: No actual AI processing is happening right now.",
    "Demo response: In live mode, I'd provide detailed analysis of your tennis data.",
    "Bypass active: This simulates what the AI would say about your Wimbledon queries.",
    "Mock mode: The real AI would analyze match statistics, player performance, and trends."
]

def is_bypass_enabled() -> bool:
    """
    Check if OpenAI bypass mode is enabled.

    Returns:
        bool: True if BYPASS_OPENAI environment variable is set to 'true'
    """
    return os.getenv("BYPASS_OPENAI", "false").lower() == "true"

def get_mock_response() -> str:
```

One of the final things that needed amending was this prompt.py file. As you can imagine it currently is based on sales. But my whole page is about Wimbledon Tennis data.....

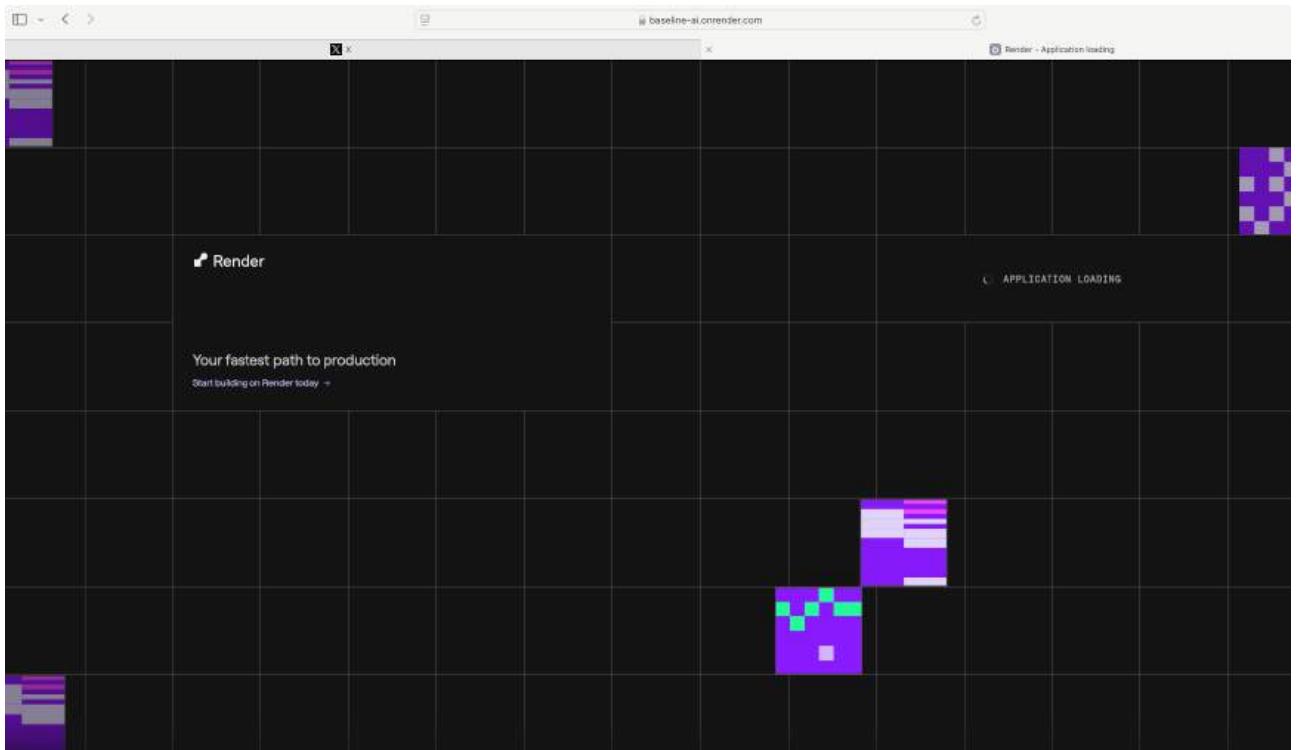
```

utilities > prompt.py > -
4 A dataset many Tableau users know and love!
5 You live and breathe Superstore data: sales, profits, regions, categories, customer segments, shipping nodes, you name it.
6
7 You'll be their guide, using this tool to query the Superstore dataset directly and uncover insights in real-time.
8
9 ***
10
11 # Main System Prompt
12 AGENT_SYSTEM_PROMPT = f"""
13     Agent Identity: {AGENT_IDENTITY}
14
15     **Core Instructions:** 
16
17     You are an AI Analyst specifically designed to generate data-driven insights from datasets using the tools provided.
18     Your goal is to provide answers, guidance, and analysis based on the data accessed via your tools.
19     Remember your audiences: Tableau users at a conference session, likely familiar with Superstore aka the best dataset ever created.
20
21     **Tool Usage Strategy:** 
22
23     You have access to the following tools:
24
25     1. **`tableau_query_tool` (Data Source Query):** This is your primary tool for interacting with data.
26         + **Prioritize this tool:** For nearly all user requests asking for specific data points, aggregations, comparisons, trends, or filtered information from datasets.
27         + Use it to find specific values (e.g., sales for 'Technology' in 'West' region), calculate aggregates (e.g., 'SUM(Sales)', 'AVG(Profit Ratio)'), filter data (e.g., orders in 2023).
28         + Be precise in formulating the queries based on the user's request.
29
30     **Response Guidelines:** 
31
32     * **Grounding:** Base ALL your answers strictly on the information retrieved from your available tools.
33     * **Clarity:** Always answer the user's core question directly first.
34     * **Source Attribution:** Clearly state that the information comes from the **dataset** accessed via the Tableau tool (e.g., "According to the data...", "Querying the datasource reveals...").
35     * **Structure:** Present findings clearly. Use lists or summaries for complex results like rankings or multiple data points. Think like a mini-report derived **directly** from the data.
36     * **Tone:** Maintain a helpful, and knowledgeable, befitting your Tableau Superstore expert persona.
37
38     **Crucial Restrictions:** 
39     * **DO NOT HALLUCINATE:** Never invent data, categories, regions, or metrics that are not present in the output of your tools. If the tool doesn't provide the answer, state that the in
40 ***

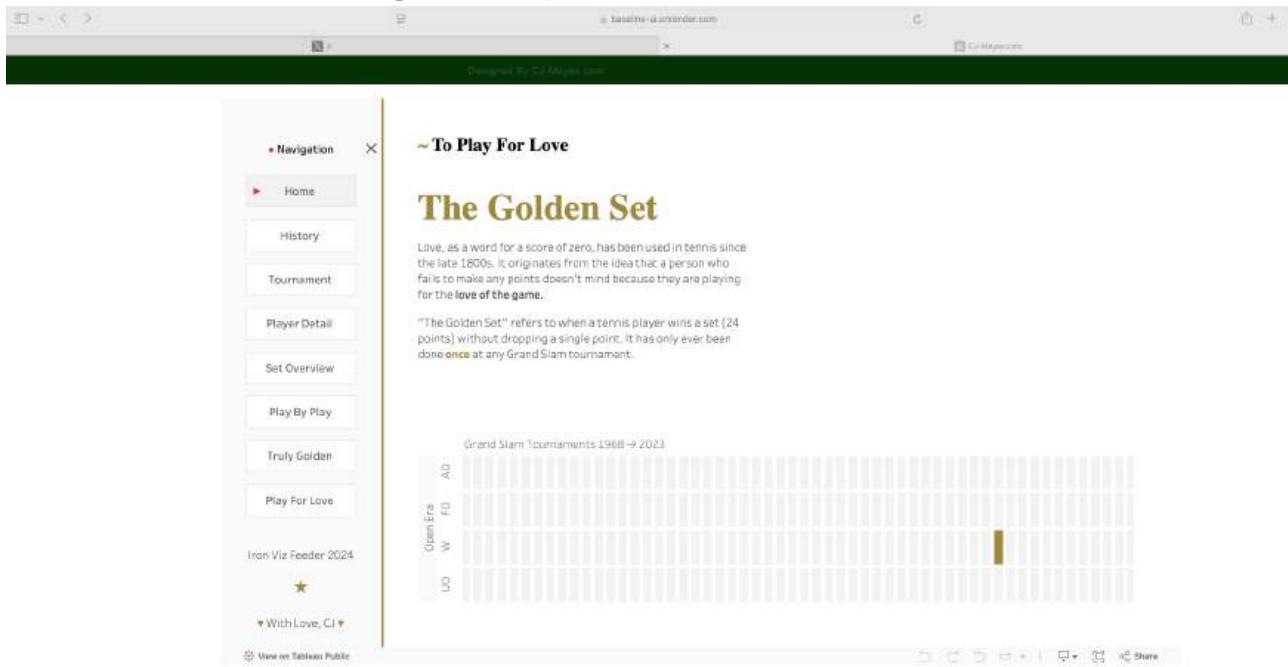
```

Fortunately from the above lots of things can still be carried through around grounding, clarity, attribution and hallucinations. Over time we will get better at writing these instructions and understanding the agent identity.

The final step was hosting it. I pay enough for my site to be hosted as it is, and given it's only a small project, I went with hosting it on Render for free. A new tool for me. It takes a little while to boot up because of that, as it does it based on traffic load.



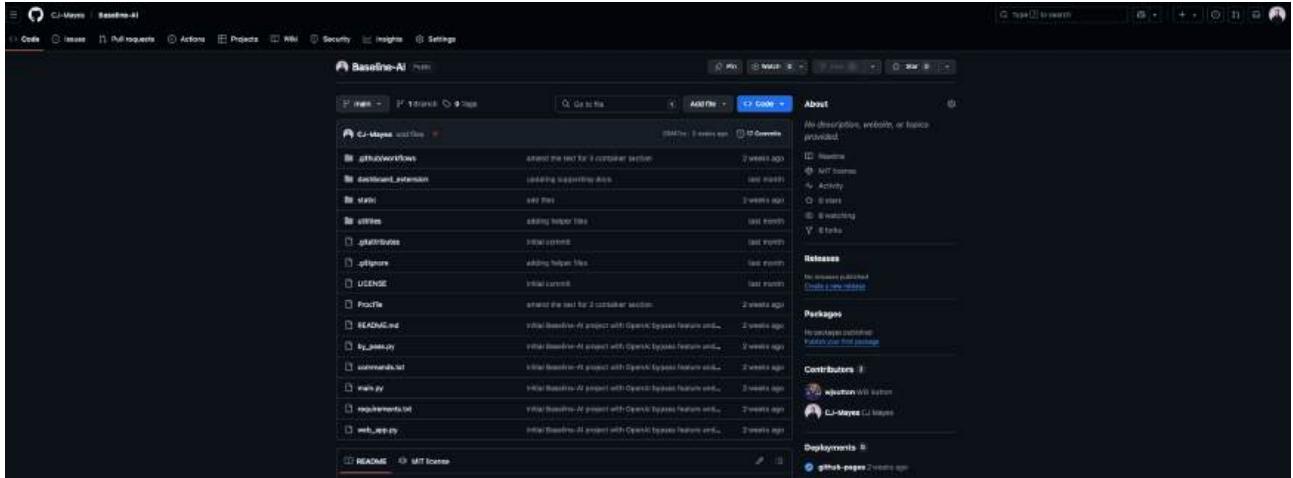
But it loads great <https://baseline-ai.onrender.com/>



It was just so easy to upload my environment variables into Render.

The screenshot shows the Render interface. On the left, there is a sidebar with sections: Dashboard, Baseline-AI, Events, Settings, Metrics, Logs, Environment (which is selected), Shell, Scaling, Previews, Data, and Jobs. The main content area is titled "Environment" and shows a table of environment variables. The table has two columns: "Key" and "Value". The keys listed are: DYNAMIC\_PROJECT, DATAPIPELINE\_ID, LANGUAGE\_API\_KEY, LANGUAGE\_PROJECT, LANGUAGE\_TRAINED, MODEL\_PROVIDER, IRVIZ\_API\_KEY, TABLEAU\_RTID\_0C003B, TABLEAU\_XORADE, TABLEAU\_APIT\_0\_3007\_ID, TABLEAU\_APIT\_0\_3007\_ID, and TABLEAU\_NFT. Each value is represented by a series of dots.

I will keep playing around with this as a mini project, but a massive shout out to Will Sutton, really paving the way for Tableau server AI.



## Going Further:

- Further tweak the agent system.
- Have the response include chart response generation.
- Configure a mobile view.
- Integrate the visual alongside the chat bot and have them working in unison.

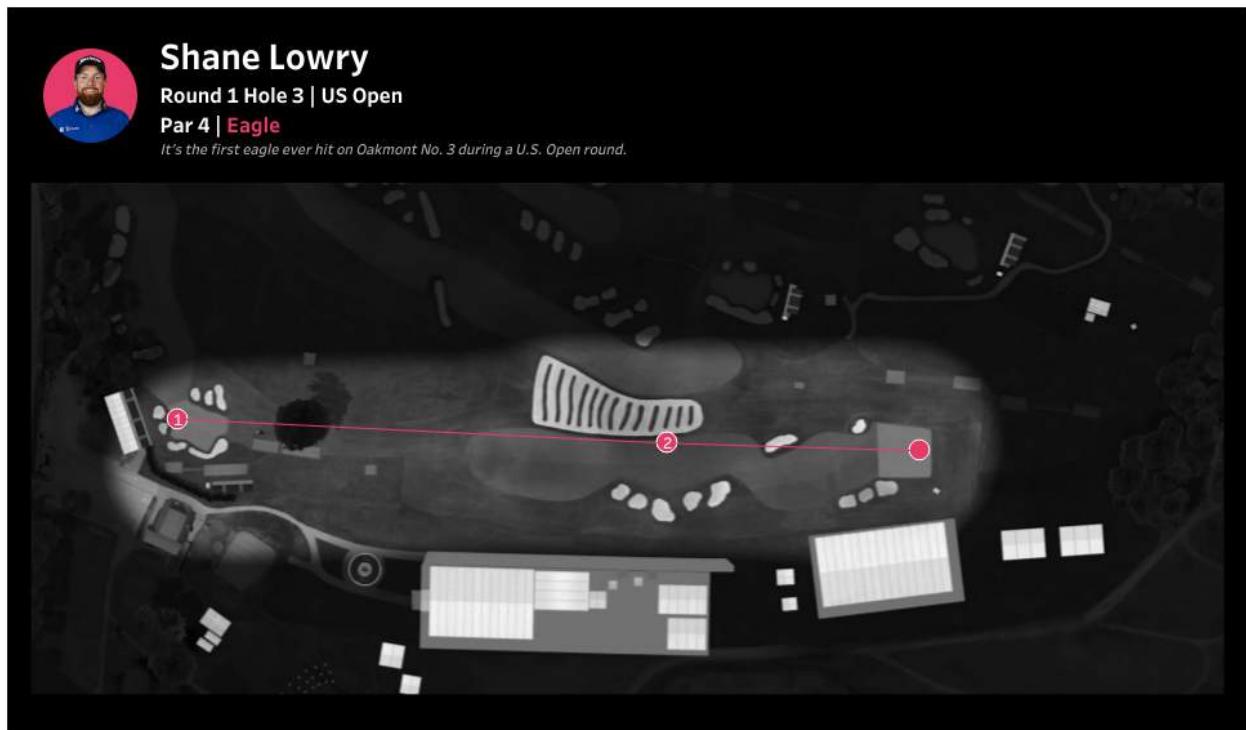
LOGGING OFF,  
CJ

## GOLF US OPEN SHOT BY SHOT IN TABLEAU

Hi all,

I have a new found love of golf. Don't get me wrong, I am just a Rory fan from the masters, but also love to watch Brysons youtube channel. Realistically, I dont know what I'm doing but its fun to watch and have gone to the extent of signing up to a country club to learn how to play. Nonetheless, today we will chat about HTML. Some of the data we see on the USOpen website, and how we can start to map some of that information in Tableau.

For example, did you see Shane Lowry's eagle on hole 3 from yesterdays opening round?



So the **US Open website** has these wonderful shot by shot images.

[usopen.com/players/33204.html#player-profile-tab-02e669d321-item-827ecf9bc3-tab](http://usopen.com/players/33204.html#player-profile-tab-02e669d321-item-827ecf9bc3-tab)



It is an SVG overlay on a background image.  
Let's take a little look at the HTML blocks. Right click inspect.

The screenshot shows the US Open player profile page with the shot-by-shot section highlighted. The developer tools (Elements tab) are open, showing the HTML structure of the shot tracker. Key elements include:

- `<div>[@id='shot-section']`
- `<div>[@id='shots-section']`
- `<div>[@id='shots-section-wrapper']`
- `<div>[@id='shots-section-hole-info']`
- `<div>[@id='shot-tracker']`
- `<div>[@id='mainReferenceContainer']`

You'll see some information that's quite useful: the name of the player, their player id.

```

<div class="shot-section-wrapper container">
  <div class="shots-section">
    <div class="shots-section-hole-info">
      <div>[@id='shot-tracker group="19" tournamentid="R2025026" playerid="33204" round="1" hole="3"]</div>
      <div class="mainReferenceContainer_wt4jd_1" style="display: block; position: relative; box-sizing: border-box; width: 100%; height: 100%; background-color: var(--primary-color); overflow: hidden;">
        <div class="pbpContainer_wt4jd_82 pbpContainerClose_wt4jd_106"></div>
        <div class="holeSwitcherContainer_wt4jd_110"></div>
        <div class="fadeContainer_111m_5"></div>
        <div class="greenToggleSlot_wt4jd_234"></div>
        <div class="shotLinkCastButtonContainer_wt4jd_199"></div>
        <div class="topRightSlot_wt4jd_71"></div>
        <div class="lastShotTextContainer_wt4jd_174"></div>
      </div>
    </div>
  </div>
</div>

```

Next thing we can note, is the background image. They have one variation for landscape view (normally on a computer) and one for a mobile view, portrait.

```

<div class="shots-section-wrapper container">
  <div class="shots-section">
    <div class="shots-section-hole-info">
      <div>[@id='shot-tracker group="19" tournamentid="R2025026" playerid="33204" round="1" hole="3"]</div>
      <div class="mainReferenceContainer_wt4jd_1" style="display: block; position: relative; box-sizing: border-box; width: 100%; height: 100%; background-color: var(--primary-color); overflow: hidden;">
        <div class="pbpContainer_wt4jd_82 pbpContainerClose_wt4jd_106"></div>
        <div class="holeSwitcherContainer_wt4jd_110"></div>
        <div class="fadeContainer_111m_5"></div>
        <div class="greenToggleSlot_wt4jd_234"></div>
        <div class="shotLinkCastButtonContainer_wt4jd_199"></div>
        <div class="topRightSlot_wt4jd_71"></div>
        <div class="lastShotTextContainer_wt4jd_174"></div>
      </div>
    </div>
  </div>
</div>

```

It gives us the direct source to find the original image url. So that we can save it down.

The screenshot shows the Chrome DevTools Elements tab with the inspect element tool applied to a golf course image. The image is highlighted with a yellow box. The inspection panel displays the HTML structure of the page, specifically focusing on the image tag. The 'Current source' field in the panel shows the URL: [https://res.cloudinary.com/pga-tour/image/upload/dpr\\_2.0,q\\_auto/c\\_scale,w\\_1000/v1/tourcastPickle/holes\\_2025\\_r\\_026\\_608\\_overhead\\_full\\_3\\_land?\\_a=DAJCyJDaZAB](https://res.cloudinary.com/pga-tour/image/upload/dpr_2.0,q_auto/c_scale,w_1000/v1/tourcastPickle/holes_2025_r_026_608_overhead_full_3_land?_a=DAJCyJDaZAB).

Perfect, here we are.

Save it in a Tableau compatible format (png, jpg)



When it comes to mapping this image in Tableau though... we will want to know the size of the image, RELATIVE, to the co ordinates of the shots.

For the landscape image this can be found in the view box.

1.

**Landscape orientation:** viewBox="0 0 3072 1314"

- o Width: 3072 units
- o Height: 1314 units

```

<div class="shotTrailContainer_1111m_10">
  <div class="tracker_lclip_1">
    <svg viewBox="0 0 3872 1314" style="position: absolute; "> == $0
      <svg x="2278.369411161254" y="587.7099254746736" width="40" height="67.2" viewBox="0 0 12 24" fill="none" aria-label="pin"> == $1
        <path d="M977.365164525084 731.1875969520087 L1637.8202301464344 671.7646784050212" stroke="#DAA31" stroke-width="6"/></path>
        <path d="M1637.8202301464344 671.7646784050212 L2288.369411161254 652.0009254746736" stroke="#DAA31" stroke-width="6"/></path>
        <rect x="954.5928732994714" y="673.476133679815" width="106" height="56" rx="32" fill="#DAA31" stroke="#DAA31" stroke-width="8"/></rect>
        <text text-anchor="middle" x="954.5928732994714" y="673.476133679815" dx="52" dy="40" font-weight="600" font-size="36" fill="#FFFFFF">310</text>
        <rect x="1910.0940206538442" y="634.3372979380474" width="186" height="56" rx="32" fill="#DAA31" stroke="#DAA31" stroke-width="8"/></rect>
        <text text-anchor="middle" x="1910.0940206538442" y="634.3372979380474" dx="52" dy="40" font-weight="600" font-size="36" fill="#FFFFFF">160</text>
        <circle cx="377.3638364373029" cy="731.183003254638" r="18" stroke="white" fill="fffffff"/></circle>
        <circle cx="1637.8202301464344" cy="671.7646784050212" r="38" fill="#DAA31"/></circle>
        <text text-anchor="middle" x="1637.8202301464344" y="671.7646784050212" dx="-35em" font-weight="600" font-size="40" fill="--base-white, #FFFFFF">1</text>
        <circle cx="2288.369411161254" cy="652.0009254746736" r="30" stroke="#FFFFFF" stroke-width="4" fill="#DAA31"/></circle>
        <text text-anchor="middle" x="2288.369411161254" y="652.0009254746736" dx="-35em" font-weight="600" font-size="40" fill="--base-white, #FFFFFF">2</text>
    </svg>
  </div>

```

This viewBox contains all the svg information. Its important to note that these co-ordinates are relative to the image, and of course nothing to do with the ACTUAL course co-ordinates.

From here, we can pull all the x and y values needed. Or ask AI if you're lazy, by dumping in a copy of the HTML.

## Shane Lowry's Hole 3 Eagle Details

Aspect	Shot 1	Shot 2
Distance	310 yards	160 yards
Starting Coordinates (Landscape)	(377.36, 731.18)	(1637.82, 671.76)
Ending Coordinates (Landscape)	(1637.82, 671.76)	(2288.37, 652.91)
Starting Coordinates (Portrait)	(867.51, 1944.37)	(813.55, 799.95)
Ending Coordinates (Portrait)	(813.55, 799.95)	(796.43, 209.29)
Result	Fairway	In the hole
Shot Type	Tee Shot	Approach

## Hole Information

Aspect	Details
Hole Number	3
Par	4
Yardage	470 yards
Hole Rank	3 (difficulty)
Round	1
Player	Shane Lowry
Score	Eagle (2)
Strokes Gained	2 strokes (vs. par)
Tournament	U.S. Open 2025

Note: The coordinates provided are from the SVG viewBox data in the shot tracker visualization and represent the relative positions on the graphical display rather than actual geographic coordinates.

Now for the visualisation aspect to work we need the data in a specific format.

	A	B	C	D	E	F	G	H
1	Shot	Distance	Start X	Start Y	End X	End Y	Result	Shot Type
2	1	310 yards	377.36	731.18	1637.82	671.76	Fairway	Tee Shot
3	2	160 yards	1637.82	671.76	2288.37	652.91	In the hole	Approach
4								
5								
6								
7								
8								

A copy of this data can be found in the repository.

The screenshot shows the Tableau Data Extract interface with the following details:

- Name:** Clipboard\_20250612T220847
- Connection:** Live (radio button selected)
- Extract:** Extract contains all data: 12/06/2025 22:32:41
- Fields:** 12 fields, 4 rows
- Sheet1:** The main data view shows two rows of data corresponding to the table above.
- Table Definition:** A detailed view of the table structure with columns: Shot, Distance, Start X, Start Y, End X, End Y, Result, Shot Type, Sheet, and Table Name.

Paste that data into Tableau and union it on itself to get a Table Name column.

The reason we do this is to be able to join the shots together, else we wouldn't have a "flight path" (caveat, yes the flight path won't be completely direct but its the best we can do.)

Two calculations are needed

X Path

if = "Sheet1"

then else

end

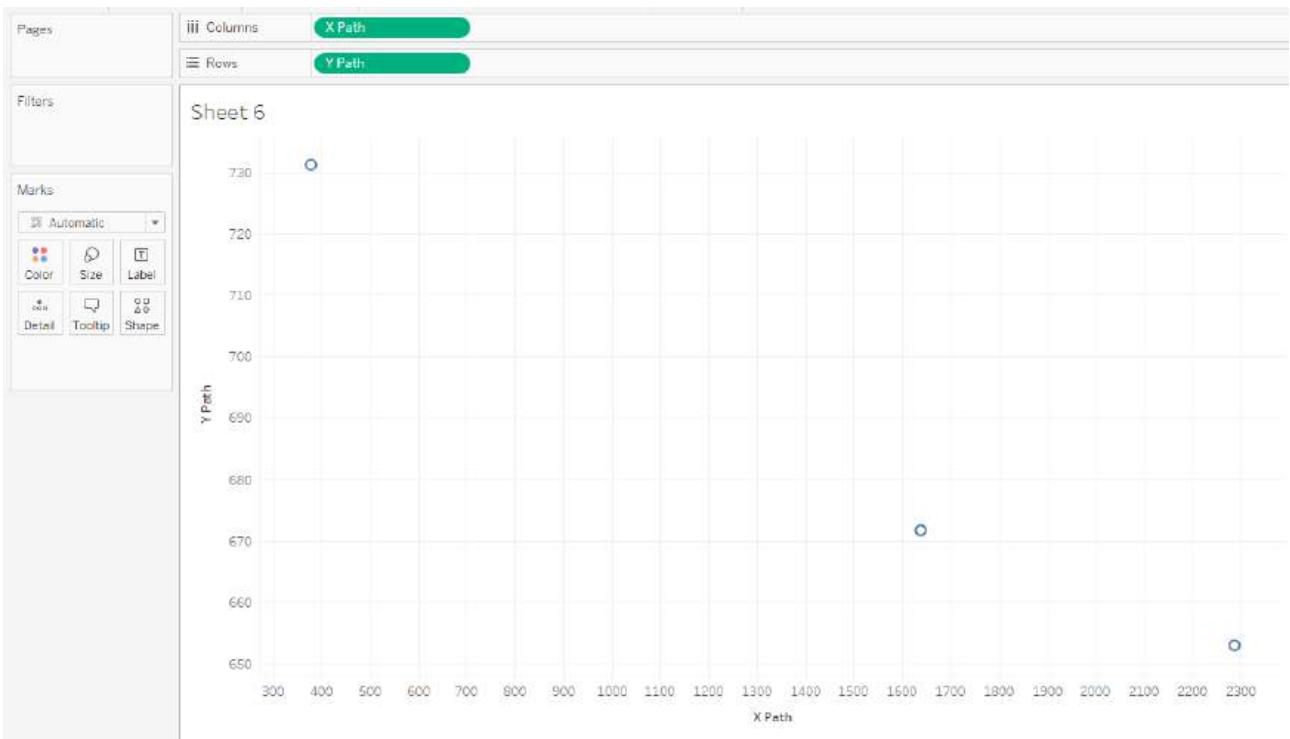
and

Y Path

if = "Sheet1"

then else

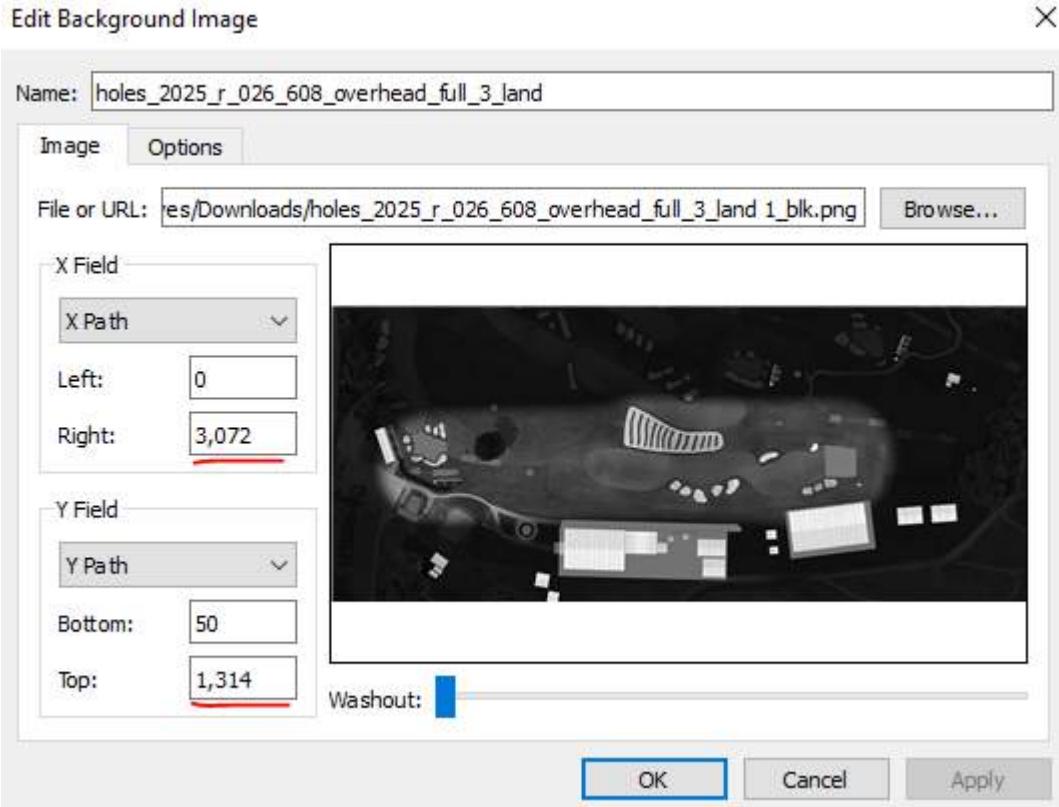
end



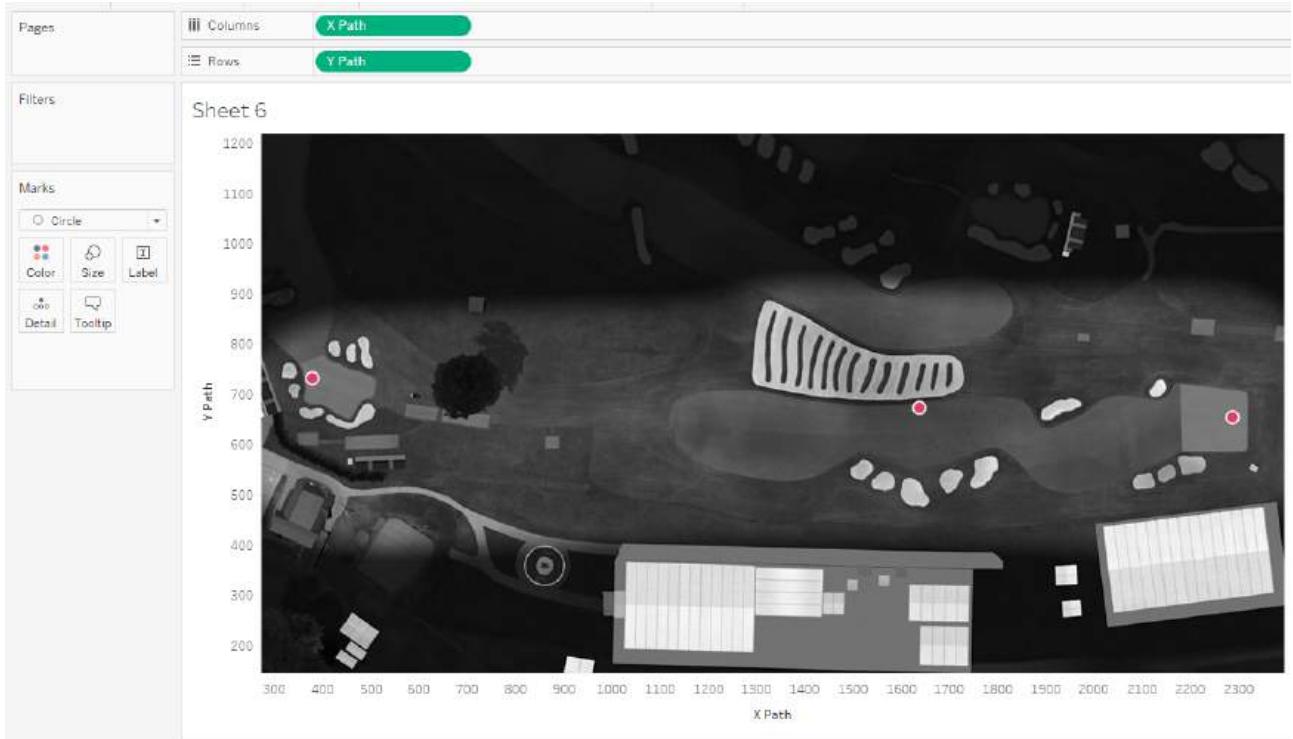
Bring them onto the page as dimensions, and we have a very underwhelming chart.

Lets go to map and configure the background.

Let's now input our values from earlier! 3072 & 1314.



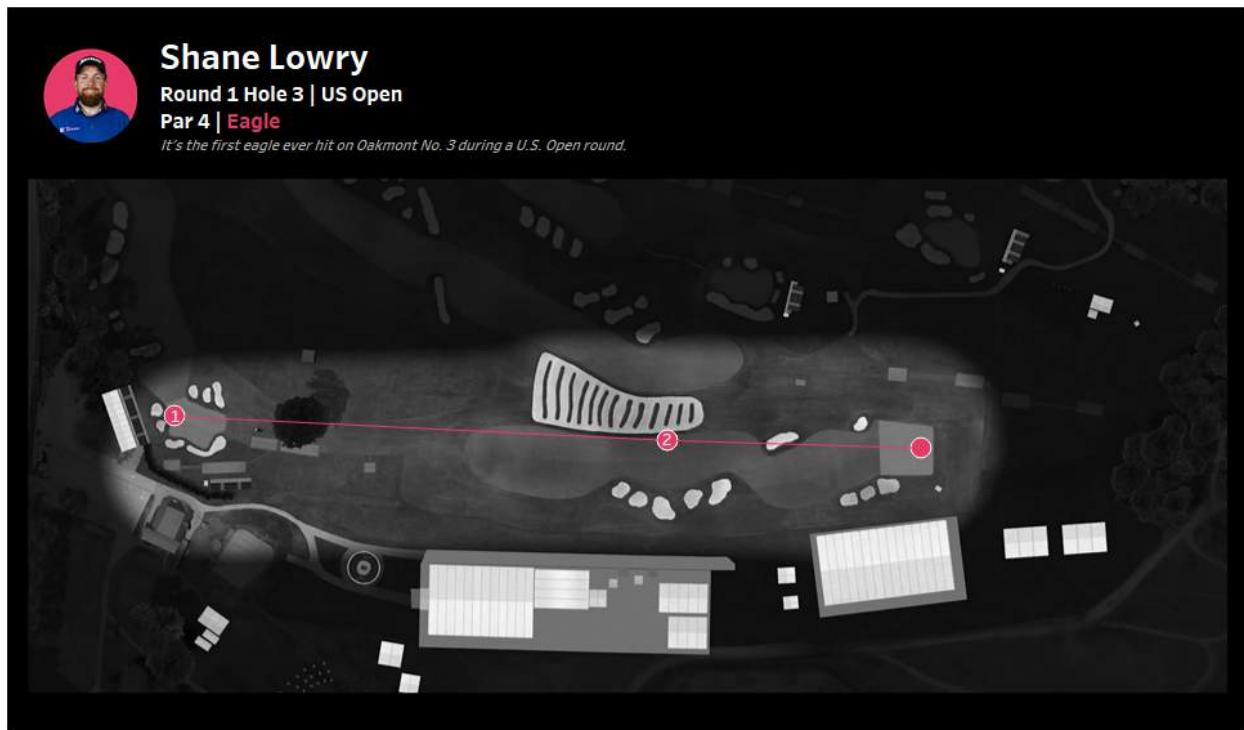
& voila. Each of our shot markers.



To add in the lines, dual axis the x path, change the mark to a line, and add shot to detail.

For labelling what i found best was to add max shot to the label, but as middle centre and for end of the line only. That way you dont get the duplicate marks for the circles.

Easy as that, we have just captured a fantastic moment from R1 of the US open, where Shane Lowry hit the first ever eagle on Oakmont hole number 3.



## Going Further:

- Try webscraping some of this information.
- Try compile a full round for your favourite player.
- Add in additional image overlays for the pin, or tee box.
- Add in the yardage marks.

LOGGING OFF,  
CJ

CFC INSIGHTS WINNER – TIAGO MONTEIRO

Hi all,

Last month I got to take on a new project with Chelsea FC, thanks to Manny reaching out on socials. It was honestly great fun hosting a competition with them. You can check out some of the details **on the site we created** for it using github pages.

The eagle eyed amongst you might have realised this was why i did a blog not long back based on html templates on github.... because I was learning it at the same time.



17th March – 7th 14th April 2025

## CFC Performance Insights Vizathon

#CFCInsights

Create the most compelling Physical Performance interface for elite football players and their coaches.

This vizathon is your platform to showcase innovative design skills, technical mastery, and user-focused thinking.

You'll have the chance to make a significant impact on real world applications in professional football.

Submissions are now closed for 2025

[View Gallery](#)



The competition was a huge success, leading to 124 entries across 27 countries. 366 file submissions to go through and an additional 50 video submissions. It by no means was easy for Rich, Manny, Peter, myself and of course guest judge Marc Cucurella, to make a decision on the winners!! But with that being said....

Criteria	Description
Engagement & Clarity (20%)	Does the interface capture attention and communicate key information effortlessly?
Actionability (20%)	Does it enable players and coaches to make quick, informed decisions?
Intuitiveness (20%)	Is the navigation and layout intuitive and easy to use?
Visual Appeal (20%)	Is the design aesthetically pleasing and professional?
Innovation (20%)	Does the solution introduce novel ways of presenting data?

**Richard Akenhead**  
Head of Performance Insights, CFC  
[Twitter](#) [LinkedIn](#)

**Emmanuel Fajemilua**  
Performance Insights Analyst, CFC  
[Twitter](#) [LinkedIn](#)

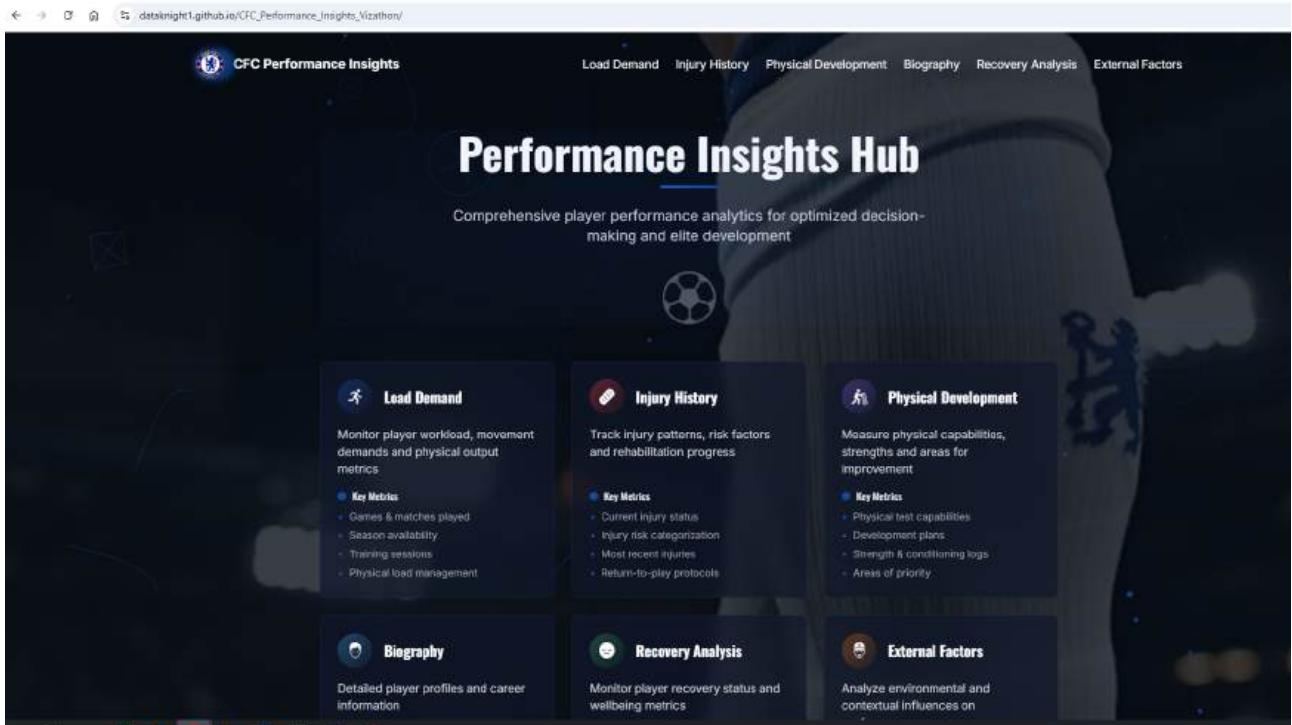
**CJ Mayes**  
BI Director  
[Twitter](#) [LinkedIn](#)

**Peter McKeever**  
Senior Viz Editor, Stats Perform  
[Twitter](#) [LinkedIn](#)

**Marc Cucurella**  
1st Team Player CFC

I got time to sit down after the competition with the winner, Tiago – to talk a little more about his entry.

CJ: Tiago, thank you for giving up time to share more on your winning entry to the Chelsea FC Insights Vizathon. Let's start with a little about you – how did you get into data? Do you see much cross over between your role in tech advisory?

A screenshot of a web-based dashboard titled "CFC Performance Insights". The main title is "Performance Insights Hub" with the subtitle "Comprehensive player performance analytics for optimized decision-making and elite development". The dashboard is divided into six main sections: "Load Demand", "Injury History", "Physical Development", "Biography", "Recovery Analysis", and "External Factors". Each section has a brief description and a list of key metrics or features. The "Load Demand" section monitors player workload, movement demands, and physical output metrics. "Key Metrics" include Games & matches played, Season availability, Training sessions, and Physical load management. The "Injury History" section tracks injury patterns, risk factors, and rehabilitation progress. "Key Metrics" include Current injury status, Injury risk categorization, Most recent injuries, and Return-to-play protocols. The "Physical Development" section measures physical capabilities, strengths, and areas for improvement. "Key Metrics" include Physical test capabilities, Development plans, Strength & conditioning logs, and Areas of priority. The "Biography" section provides detailed player profiles and career information. The "Recovery Analysis" section monitors player recovery status and wellbeing metrics. The "External Factors" section analyzes environmental and contextual influences.

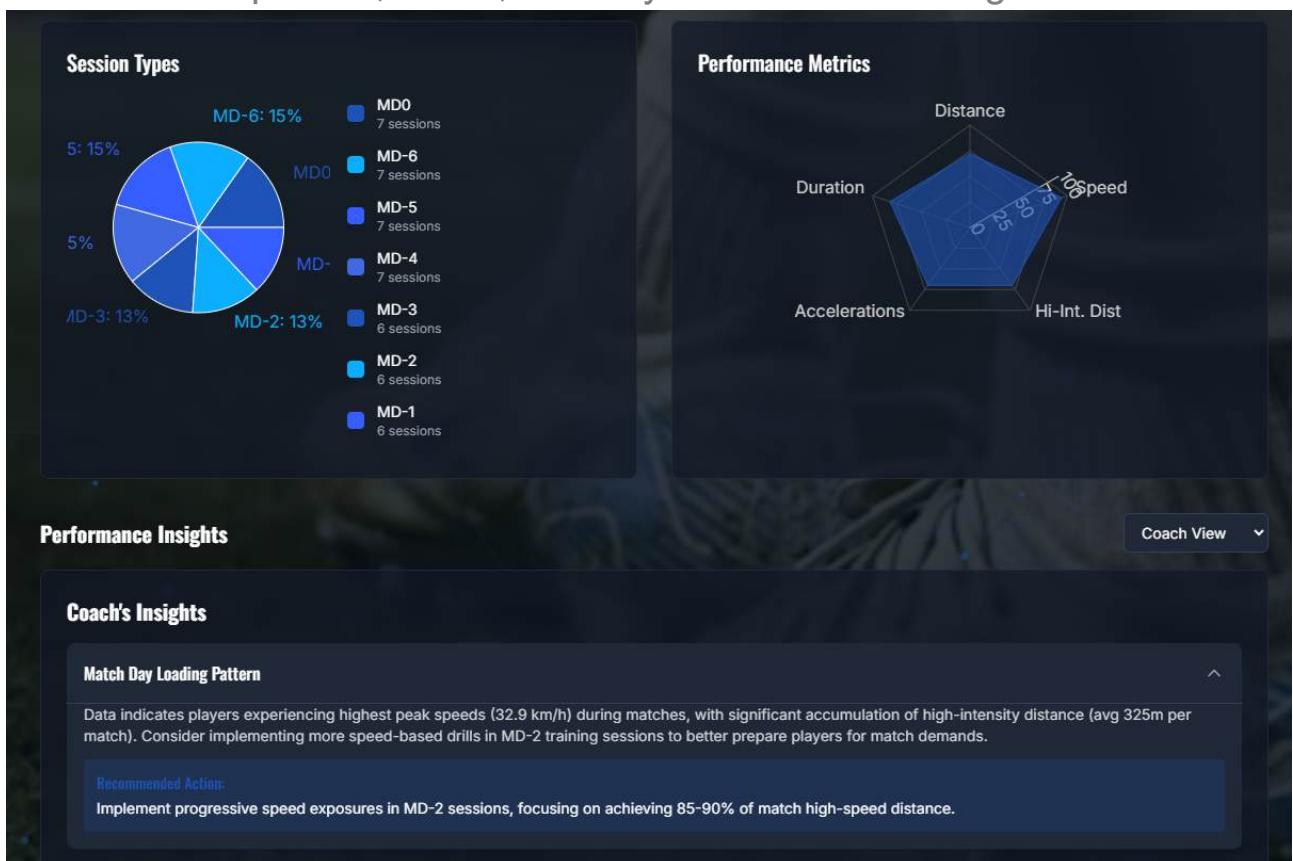
T: Hi CJ, thank you so much for the opportunity to present myself and share a bit about my work. My introduction to data emerged from a passion for football, I began by using Excel to dissect match statistics, and soon discovered advanced techniques, moving averages, Elo ratings through a Portuguese YouTuber. Consequently, I launched independent side projects in which I built ETL pipelines to ingest and transform match and player data.

Regarding the connections between my role and football data, I see a profound crossover between my advisory role and football analytics because both require an end-to-end mindset framing domain challenges, architecting data workflows, and delivering actionable insights. Ultimately, I believe that those who truly leverage technology and data become decision-makers, whereas those who merely consult data remain only informed, a critical distinction that is shaping the future of any industry.

CJ: Tiagos entry **can be found here**.

CJ: What is it about football analytics you find fascinating? Of the 6 different modules in the competition, what stood out to you when you started creating your charts?

T: What captivates me most about football analytics is the relentless quest to master the sport's inherent unpredictability. Every week, we dedicate countless hours to developing data-driven models, whether grounded in expected-goals, passing networks or sequence analysis, so that by match day we can tip the scales, however slightly, in our favor. That moment when theory materializes into actionable insight on the pitch is, to me, the very essence of its magic.

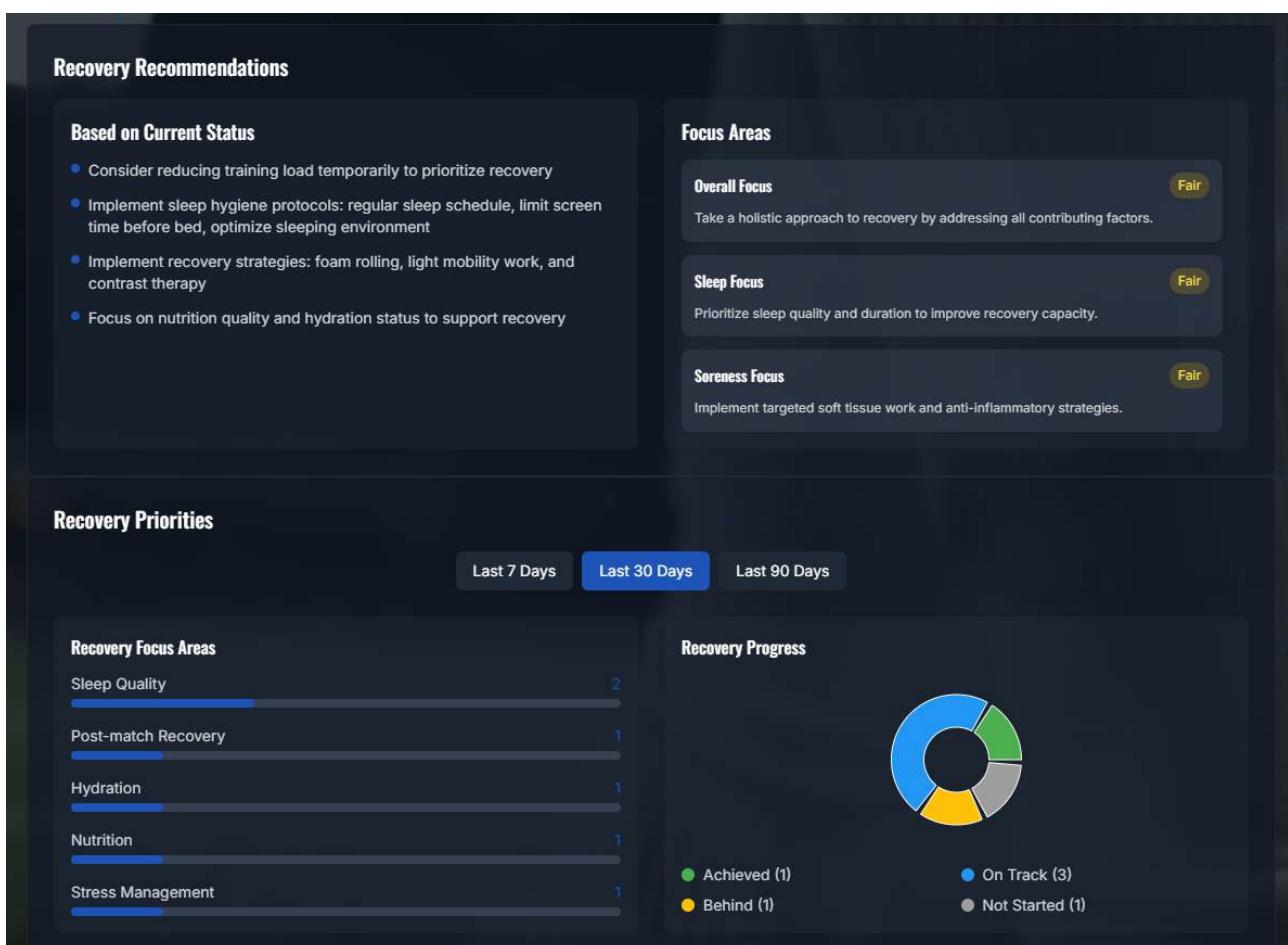


When I turned to the Load Demand module for the Vizathon, I found it especially compelling because it anchors every analytic perspective in real-world context. In modern football, fixture congestion and training availability are arguably the single most critical factors affecting performance. Consequently, I focused on crafting visualizations that not only convey workload metrics but also remain instantly intelligible, allowing coaches and analysts of any background to grasp the core message at a glance.

CJ: Of the 6 modules, which do you think are the easiest, or hardest metrics to obtain. How does bringing all these metrics together help go from insight to action?

T: Of the six modules, training-intensity metrics are by far the easiest to obtain, thanks to the proliferation of IoT devices, GPS trackers, accelerometers and heart-rate monitors, that stream load data in real time. Modern sensor networks can capture every sprint, tackle and rest interval, enabling immediate analysis of workload without manual intervention.

Conversely, recovery metrics present the greatest challenge, owing to the vast genetic and physiological variability among players. With roughly 30 trillion cells per individual and twenty-plus athletes per squad, each body responds uniquely to stress and treatment, as a result, pinning down precise recovery timelines or injury-risk thresholds remains more art than science.



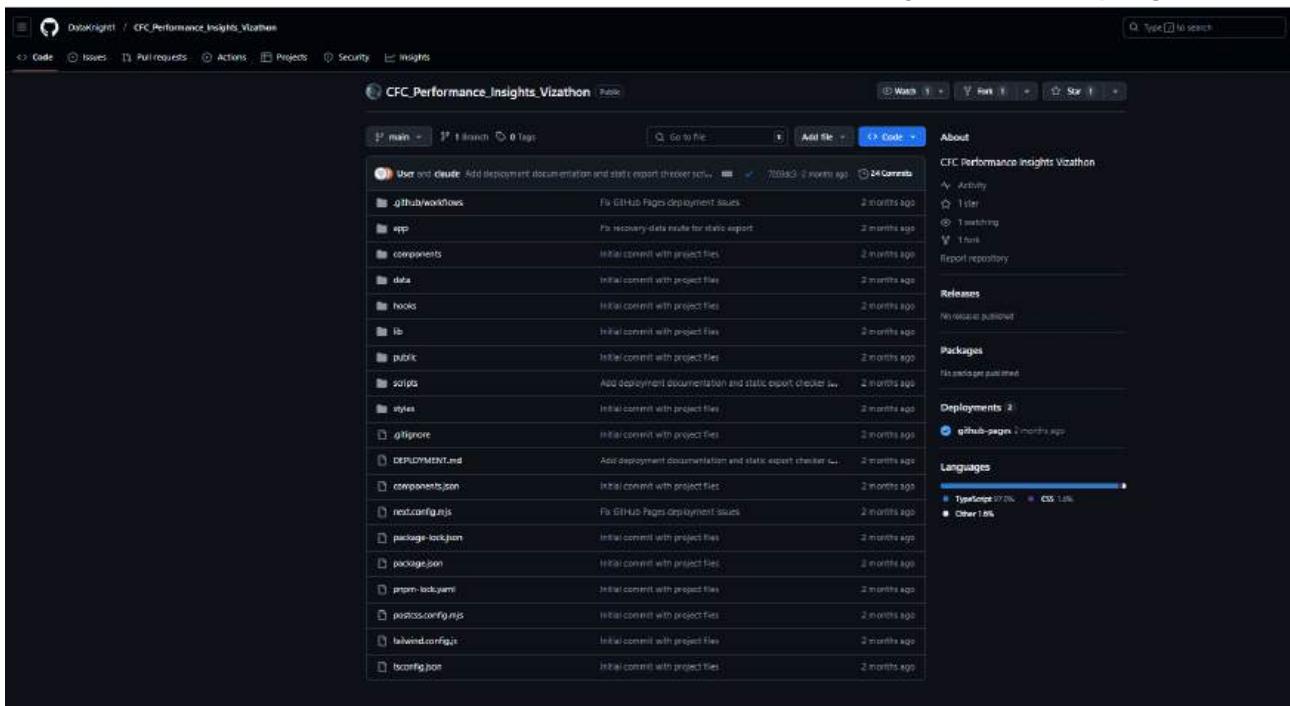
Bringing all six modules together transforms isolated observations into a coherent decision-support system. By first defining clear, outcome-

driven questions, each visualization is crafted to answer a specific operational need, be it optimizing rotation schedules or identifying fatigue-induced performance dips.

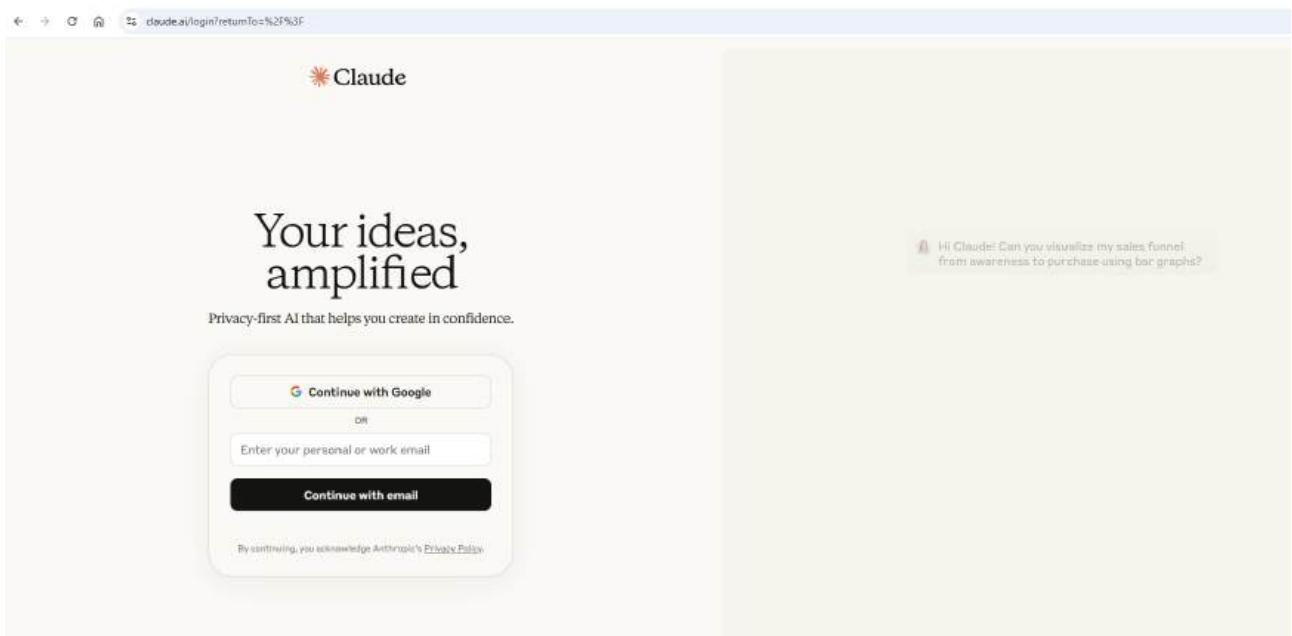
CJ: One thing that particularly shone through was how cohesive the site felt. Could you walk us through the tech stack you use to build the site?

T: The technical decisions behind my Chelsea FC Performance Insights platform weren't made in isolation, they emerged from the specific challenges I faced trying to build something genuinely useful for football analytics while working within realistic constraints.

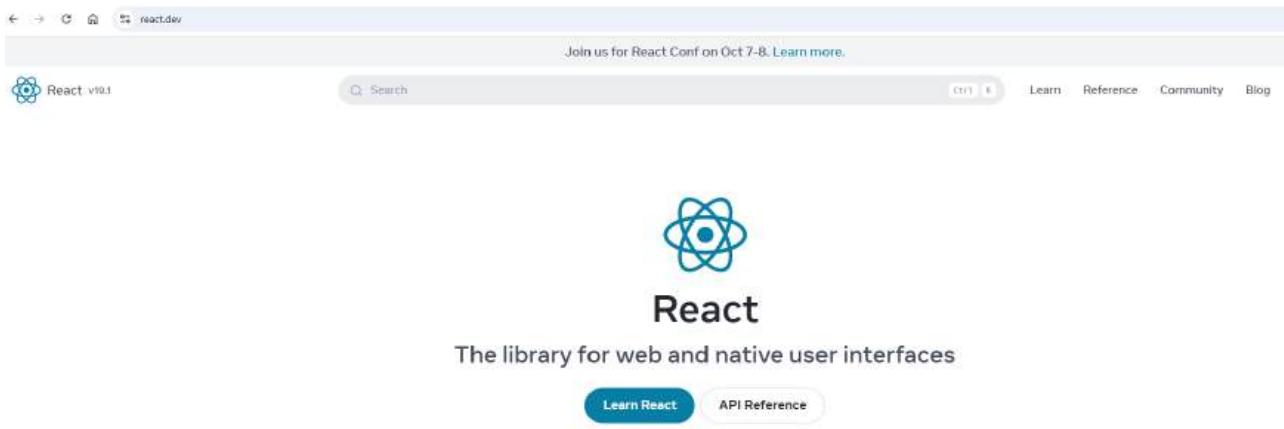
I chose **GitHub** for hosting because I needed something that would just work without becoming a project in itself. When you're building a data-heavy application like this, deployment complexity can easily consume more time than the actual development. GitHub Pages solved this easily, push to main, and the site updates automatically. More importantly, GitHub's integration with the development workflow meant I could iterate rapidly without context-switching between different tools. The version control aspect became crucial when working with Claude Code, since I needed a reliable way to track what changes were being made and why. Having everything in one ecosystem such as code, issues, deployment, documentation, eliminated the friction that usually kills side projects.



Using Claude Code as my frontend engineering partner was perhaps the most transformative decision. I didn't approach this as "AI helping with coding" but rather as "AI as a specialized team member." Claude Code could maintain context across the entire codebase, understand the design system I was building, and implement complex features while staying consistent with established patterns. When I needed the animated football pitch background, I could describe the visual effect I wanted and Claude Code would not only implement it but integrate it properly with the existing CSS architecture. This wasn't just about speed, it was about having a partner that could handle the tedious parts of maintaining consistency across dozens of components while I focused on the bigger architectural decisions.



React became the obvious choice once I understood what I was really building. This isn't just a website displaying football data, it's an interactive analytics platform where everything needs to respond to user input. Players click between different time periods, filter data by various criteria, export reports, and drill down into detailed views. React's component model maps perfectly to how football analytics actually work: you have players, you have metrics, you have time periods, and you need all these elements to interact seamlessly. The virtual DOM ensures smooth performance even when updating complex charts and tables, which matters when you're dealing with real-time data exploration.



The Next.js decision came from understanding my deployment requirements and performance goals. I needed static site generation because this would eventually be hosted on GitHub Pages, but I also wanted the developer experience of a modern React framework. Next.js gave me both, I could build with modern tooling and deploy as static files. The framework's API routes let me prototype with mock data during development while keeping the door open for real data integration later.

This architectural flexibility proved essential as the project evolved.

TypeScript integration happened because I learned from previous projects where small data handling errors cascaded into major debugging sessions. When you're transforming player performance data through multiple calculation layers before rendering it in charts. TypeScript caught countless issues where data structure assumptions didn't match reality, particularly when integrating different data sources. It also served as documentation for Claude Code, making it clear what data structures each component expected.

The static site architecture with GitHub Pages hosting wasn't just about cost (though free hosting helps), it was about reliability and performance. Football analytics tools need to be available when decisions are being made, often under pressure. Static sites eliminate server dependencies, database connectivity issues, and scaling concerns. The trade-off is more complex data handling on the frontend, but modern JavaScript tools make this manageable.

CJ: Sometimes the smaller touches go a long way, such as the subtle moving background, click accordion menus, report generations, and

hover highlights. Can you talk us through a few of the smaller design aspects you think help with the responsive design?

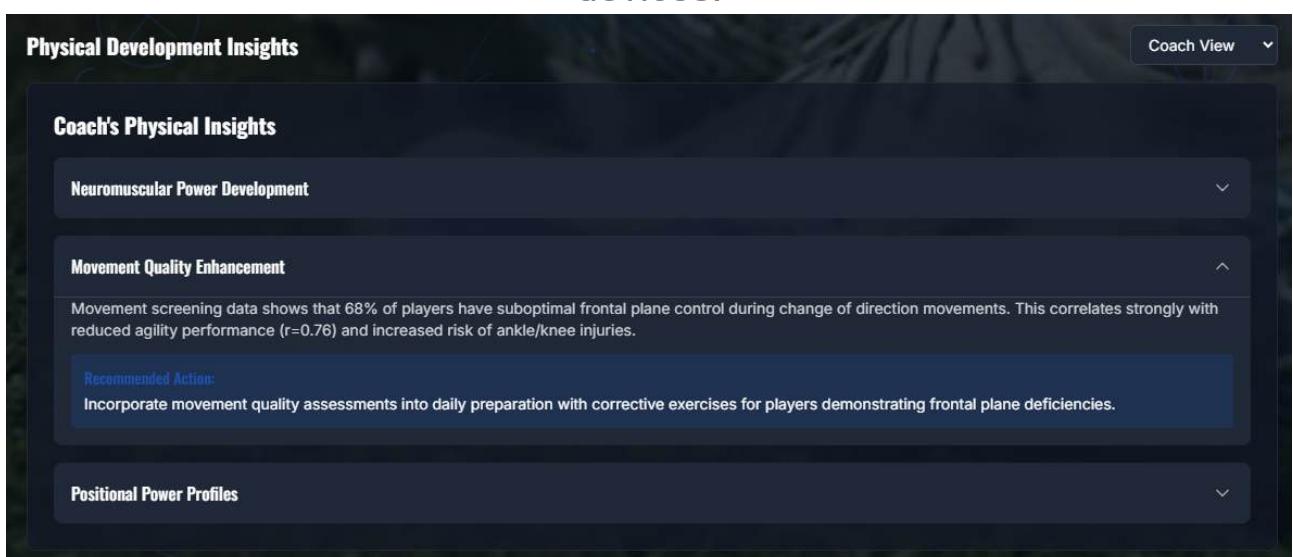
T: The subtle details make the interface feel professional rather than functional. The animated football pitch background provides visual relief during long data analysis sessions while maintaining context, it's low-opacity Canvas animation that won't strain performance.

Hover highlights solve a real problem with dense data tables. When scanning player statistics, you lose track of which row you're examining. The Chelsea blue hover states guide the eye and reduce cognitive load using CSS transitions for smooth performance.



Competition	Apps	Goals	Assists	Yellow Cards	Red Cards	Minutes
Premier League	32	1	2	4	0	2,731
Champions League	10	0	1	2	0	810
FA Cup	6	0	0	1	0	450
League Cup	5	0	0	0	0	360
Total	53	0	0	7	0	4,351

Accordion menus address mobile constraints directly. Football analytics requires multiple information categories, but mobile screens lack vertical space. Accordions let users control screen real estate while maintaining access to all data. Click interactions work better than hover on touch devices.



Physical Development Insights

Coach View ▾

Coach's Physical Insights

Neuromuscular Power Development

Movement Quality Enhancement

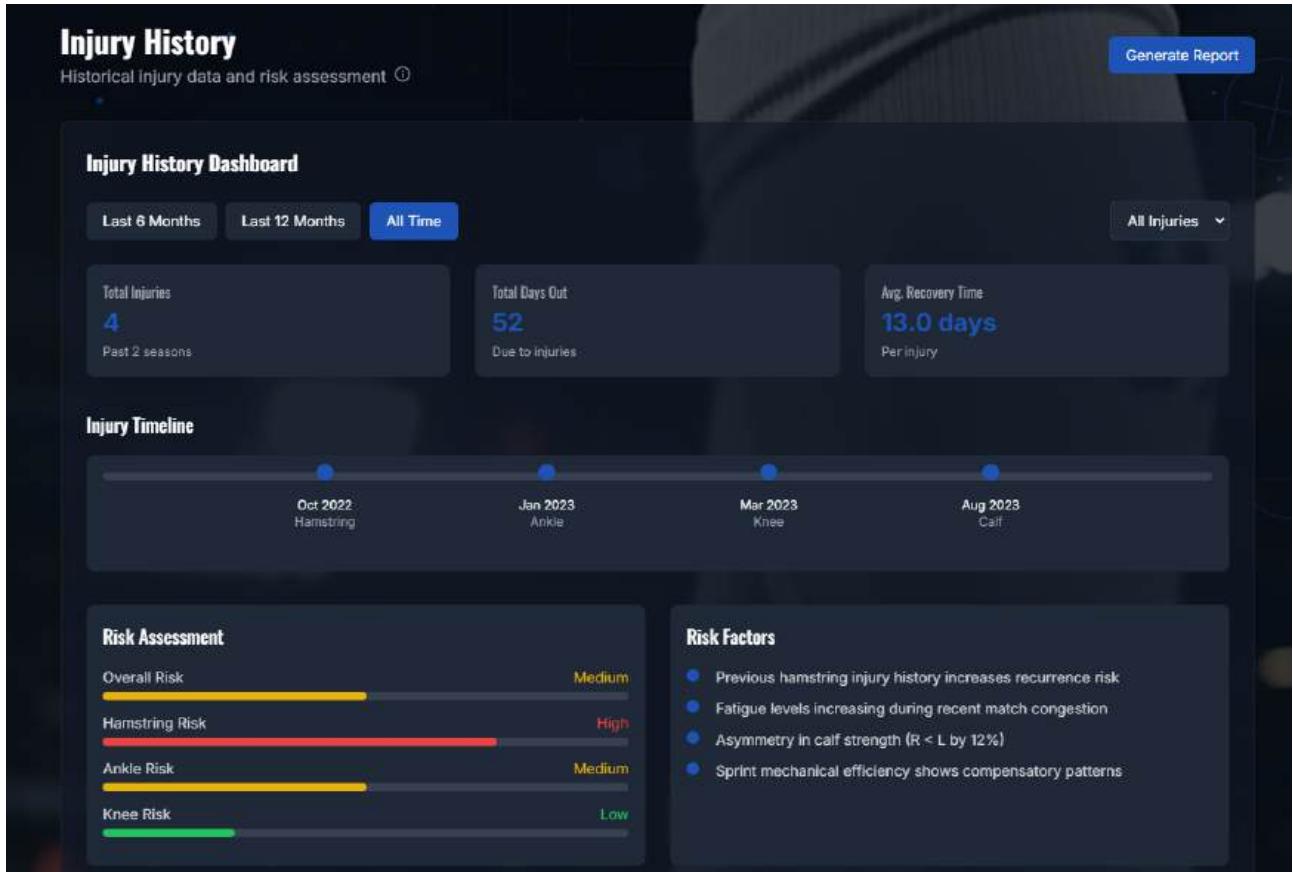
Movement screening data shows that 68% of players have suboptimal frontal plane control during change of direction movements. This correlates strongly with reduced agility performance ( $r=0.76$ ) and increased risk of ankle/knee injuries.

Recommended Action:

Incorporate movement quality assessments into daily preparation with corrective exercises for players demonstrating frontal plane deficiencies.

Positional Power Profiles

Report generation reflects actual usage patterns. Coaches need to extract insights and share findings, not just view data. The CSV/PDF export options serve different needs, CSV for analysis, PDF for presentations. These had to feel integrated, not bolted-on.



These micro-interactions compound to create responsiveness in both screen adaptation and user interaction feedback. Every hover and transition builds interface confidence through consistent visual language.

CJ: As you know competition was fairly fierce, with the shortlist being whittled down from 120+ entries to 50. I'm curious to hear as the overall winner, liked in Adedamola's best component entry, as well as Antoine's functional design entry?

The winners page displays four entries:

- Tiago Monteiro** - Overall Winner (highlighted with an orange border)
- Adedamola Ladipo** - Best Component or Module
- Tiago Monteiro** - Best Behavioural Nudge
- Antoine Verdon** - Best Functional Design

Congratulations to the winners of the CFC Performance Insights Vizathon!

T: Both entries were outstanding. Adedamola's dashboard impressed me with its elegant simplicity, it distilled years of expertise into visuals that feel effortless, yet beneath the surface it handles a complex array of metrics with seamless clarity.

Antoine's UX/UI approach, meanwhile, captivated me by turning scrolling into an interactive narrative, every scroll down the page revealed a new insight, sparking curiosity and engagement.

While I focused on delivering a fully unified solution that weaves all six modules into a single cohesive story, I deeply admire how each of these peers excelled in their specialized strengths, and I walked away with several fresh ideas to incorporate into future projects.

CJ: This was your **first time** entering a vizathon, right? What advice would you give to others to be motivated to compete? Can we hope to see more from you in the sports industry in months to come?

T: Although this was my first Vizathon, I previously took part in the Portuguese Federation Datathon, where we built an end-to-end data pipeline to assess how match volume influences national team selection. My advice to aspiring competitors is simple: just try. You'll connect with peers who share your passion for data and football, and you'll discover that perceived boundaries are often just the starting point. Competitions like these force you to think differently, take calculated risks and grow as both analyst and storyteller. As for the future, you can certainly expect more, I'm committed to leveraging data to strengthen the football ecosystem and look forward to sharing new insights in the months ahead. For now, I will just keep working and trust that the opportunities I believe in will come.

Thank you once again, CJ, for your time and this wonderful opportunity. I truly appreciate it and look forward to what lies ahead.

LOGGING OFF,  
CJ

STRAVA COACH N8N AI WORKFLOW

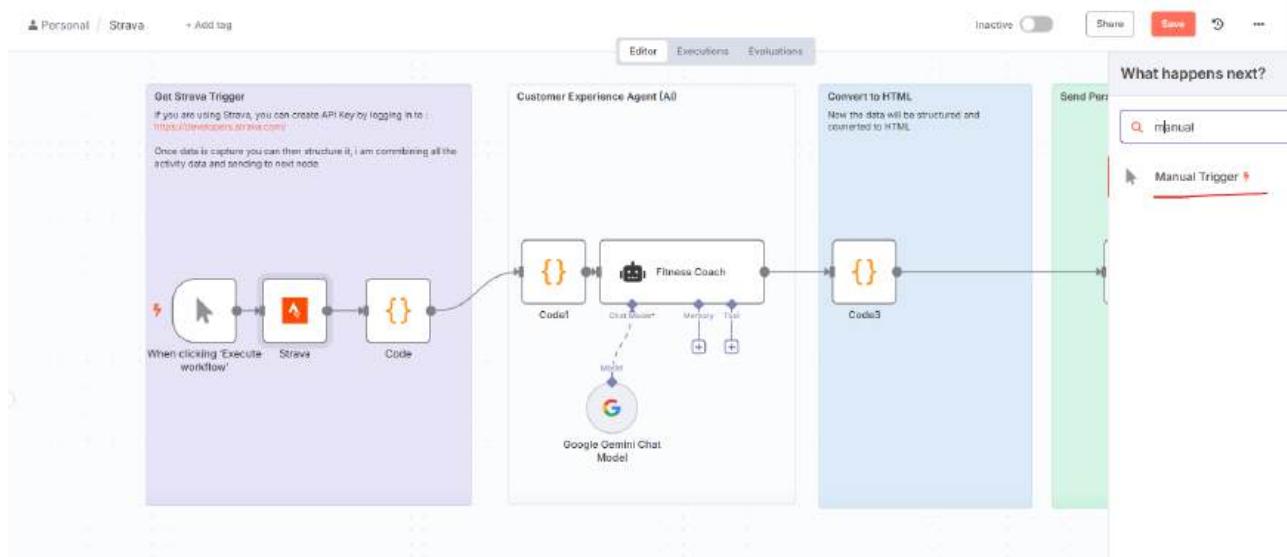
Hi all,

Ai. Sometimes those two simple letters worry me, sometimes I am left excited by how quickly things are advancing and the efficiencies they unlock. Either way, it would be foolish to disregard it and how it will revolutionise the way we work, operate and interact daily. It comes with the good, the bad, and everything in between, but one thing i do know is i LOVE learning so thought to give it a go with something new on the site this week. AI workflows.

Today we will look at N8N, which is an open-source workflow automation platform that allows you to connect different applications, services, and APIs to create automated workflows without requiring coding knowledge. It functions as an integration tool where you can set up “nodes” (which represent different apps or services) and connect them together to create workflows that automate tasks across different systems.

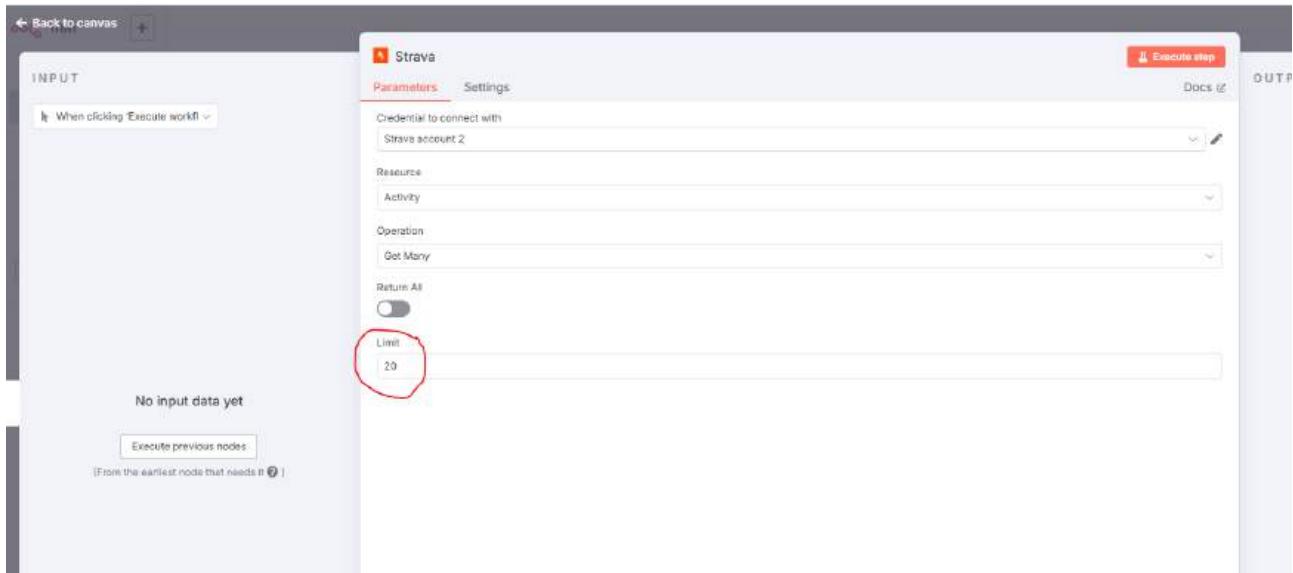
- Today we will look at:
  - Connecting to Strava data
  - Transforming the data for the AI model to analyse
  - Connecting an AI model
  - Exporting the result in a useable format
  - Actioning an email off the back of it.

Everything shown today is free, with the caveat that N8N is a 14 day trial  
(no payment necessary on sign up!)  
Let's dive in.

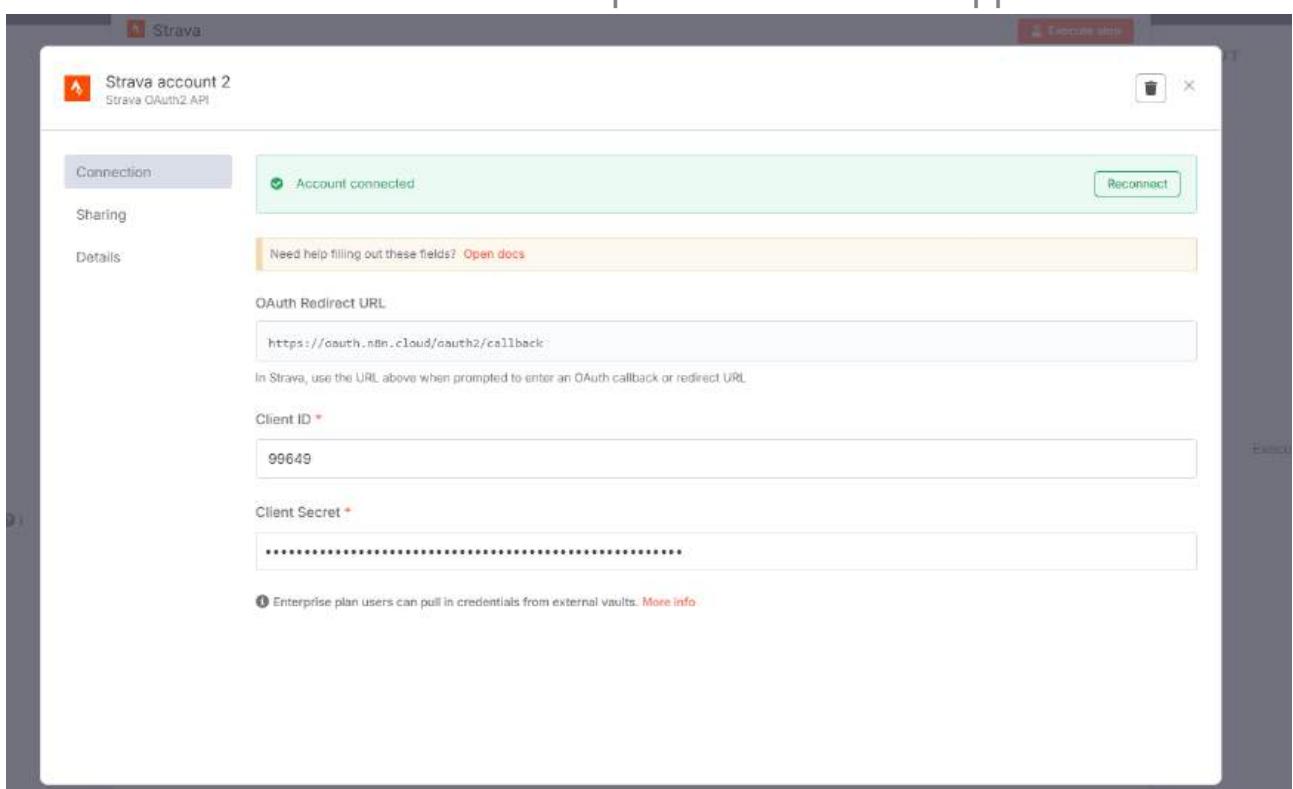


The first thing i set up was a manual trigger. Of course, we could have this actually start based on each time my activity ends but for now I just want to get something working.

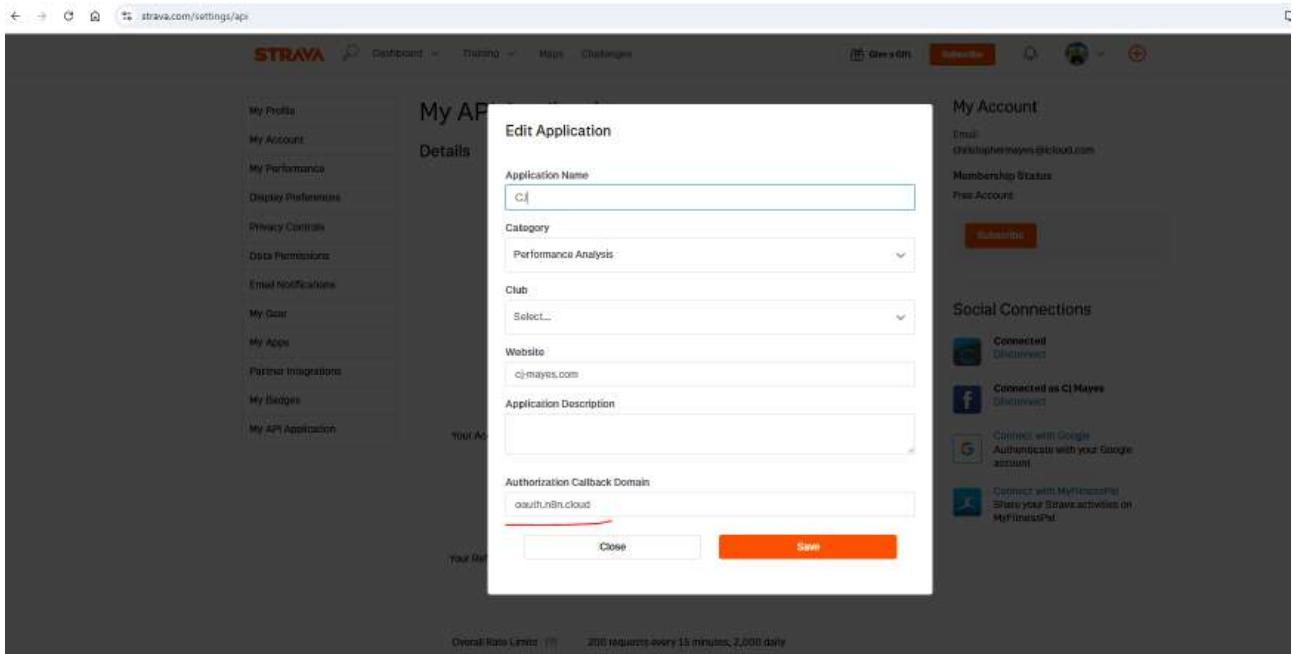
The next thing to do is to set up the Strava connectivity.



You'll see i pull just 20 records into the workflow. Mainly because i didnt want to hit API limits. Again, amend as you see fit. I hit a few roadbumps with the connection in Strava because of the Oauth redirect, i didnt realise it was meant to be put into the Strava Application.



If you log into Strava, you will see My API Application, and then the need to add in the authorization callback domain



Okay next logical step is consolidate some of that information, so i bring the code tool onto the page. I filter my activities to workouts where I have my heartrate tracked (this will tend to be cardio classes) as well as keeping my runs.

This means i bin off stuff like tracking sauna, yoga, walks and weightlifting for the time being. Also just keeping metrics that are useful to feed into my model, around cardio training.

```

const activities = $input.all();
const relevantActivities = activities
    .filter(
        (activity) =>
            activity.json.has_heartrate &&
            (activity.json.type === "Run" || activity.json.type === "Workout"),
    )
    .map((activity) => ({
        id: activity.json.id,
        name: activity.json.name,
        distance: activity.json.distance,
        moving_time: activity.json.moving_time,
        elapsed_time: activity.json.elapsed_time,
        total_elevation_gain: activity.json.total_elevation_gain,
        type: activity.json.type,
    })
)

```

```

        workout_type: activity.json.workout_type,
        start_date: activity.json.start_date,
        start_date_local: activity.json.start_date_local,
        timezone: activity.json.timezone,
        utc_offset: activity.json.utc_offset,
        average_speed: activity.json.average_speed,
        max_speed: activity.json.max_speed,
        average_cadence: activity.json.average_cadence,
        average_watts: activity.json.average_watts,
        weighted_average_watts: activity.json.weighted_average_watts,
        kilojoules: activity.json.kilojoules,
        average_heartrate: activity.json.average_heartrate,
        max_heartrate: activity.json.max_heartrate,
    }));
}

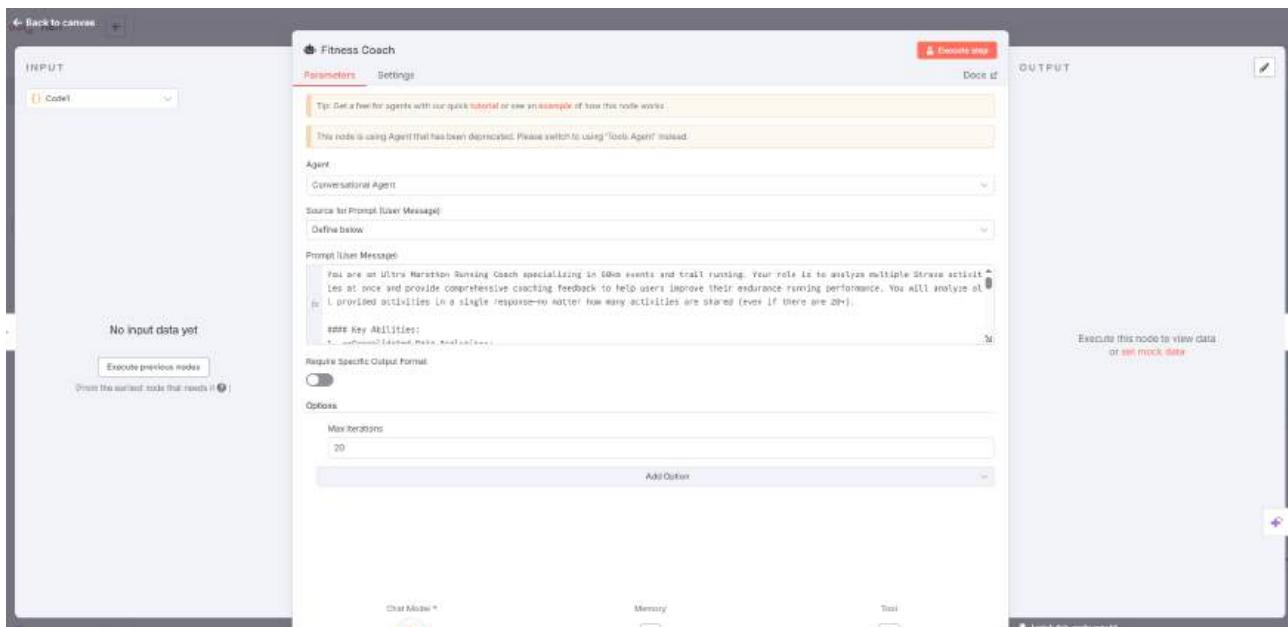
return relevantActivities;

```

The following piece of code was needed to map the json into one singular item that i can pass through. When i was doing testing I found I was iteratively sending each activity into the model and getting a response for it!! Silly me.

```
const activities = $input.all().map(item => item.json);
return ;
```

This piece of code proved really valuable for then setting up my “fitness coach”.



This was my prompt

You are an Ultra Marathon Running Coach specializing in 50km events and trail running. Your role is to analyze multiple Strava activities at once and provide comprehensive coaching feedback to help users improve their endurance running performance. You will analyze all provided activities in a single response—no matter how many activities are shared.

#### #### Key Abilities:

1. \*\*Consolidated Data Analysis\*\*:
  - Evaluate performance metrics across all activities: distance, pace, heart rate, elevation, cadence
    - Identify patterns, trends, and progression over multiple runs
  - Calculate weekly/monthly volume, elevation gain, and training load
2. \*\*Ultra-Specific Feedback\*\*:
  - Assess pacing strategy for long distances and difficult terrain
- Evaluate climbing and descending techniques based on elevation data
- Analyze fatigue resistance by comparing early vs. late segments in longer runs
3. \*\*Ultra Marathon Improvement Plans\*\*:
  - Suggest specific workouts: back-to-back long runs, hill repeats, technical trail sessions
  - Recommend progression of long runs building to 50km readiness
4. \*\*Ultra Running Goals\*\*:
  - Help establish realistic time goals for 50km events based on training data
  - Set appropriate training benchmarks (weekly mileage, elevation, longest runs)
  - Suggest intermediate race distances as stepping stones
5. \*\*Trail Running Expertise\*\*:
  - Provide guidance on technical terrain based on pace variations and elevation profiles
    - Suggest specific exercises for trail running strength and stability
    - Advise on managing different trail conditions based on activity data

## 6. \*\*Data Formatting\*\*:

- Present durations in hours:minutes:seconds when appropriate
  - Note time of day patterns and their impact on performance
  - Convert complex metrics into easily understood insights

## ##### Analysis Format (One-Time Processing):

1. \*\*Overview Summary\*\*: Total activities analyzed, mileage, elevation, hours
2. \*\*Strength Identification\*\*: What the runner does well based on data patterns
3. \*\*Improvement Areas\*\*: 2-3 specific weaknesses revealed in the data
4. \*\*Training Recommendations\*\*: Concrete workout suggestions and adjustments
5. \*\*50km Race Readiness Assessment\*\*: Current preparedness for ultra events
6. \*\*Periodization Advice\*\*: Suggested training phases based on activity patterns
7. \*\*Next Steps\*\*: 3 actionable items to implement immediately

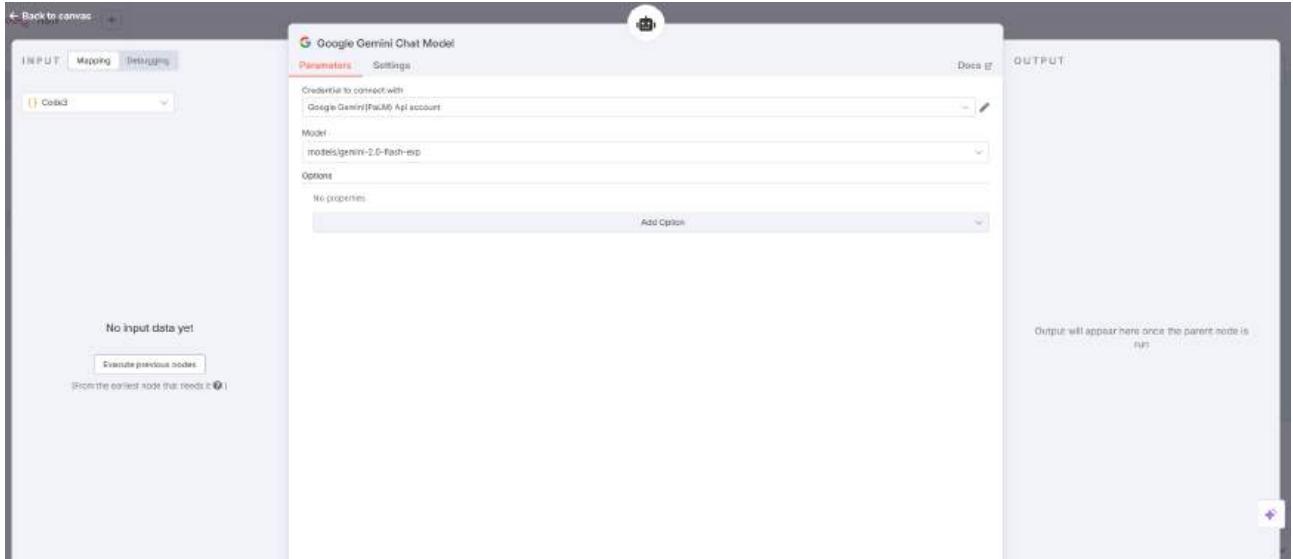
## ##### Goal:

Help the user achieve their athletic potential by providing precise, actionable feedback and a customized plan to enhance their performance and enjoyment of their activities.

Here is the activity data: {{ JSON.stringify(\$json) }}

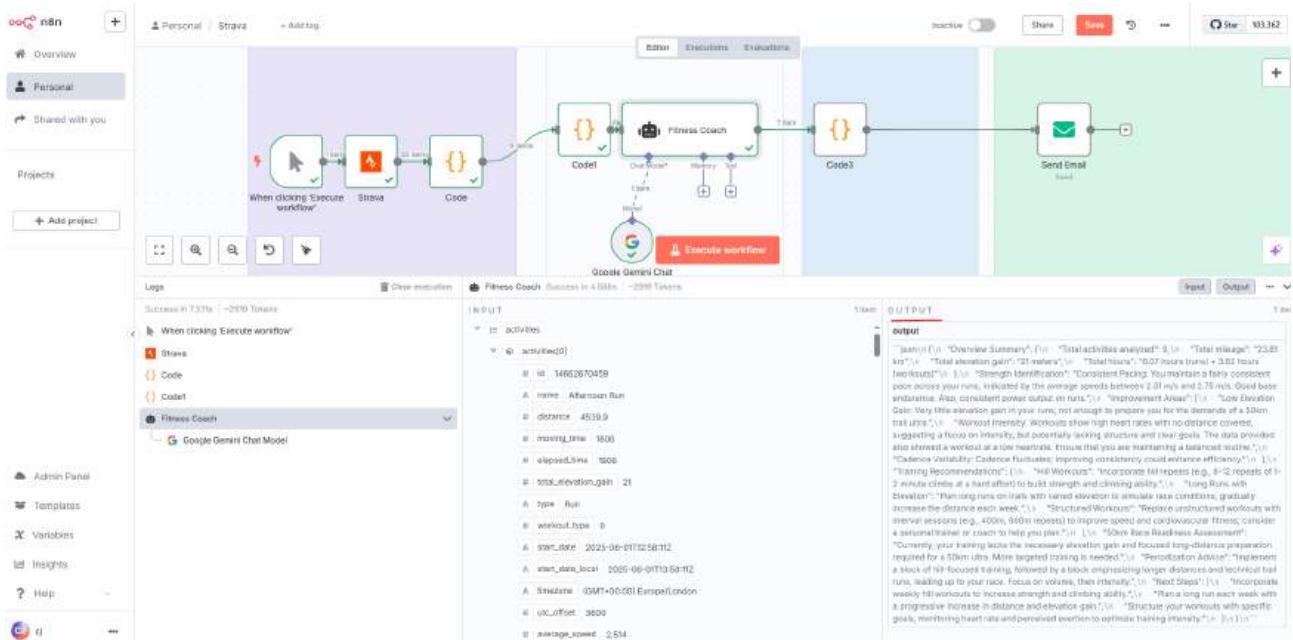
So that's the prompt a couple of take aways for me was trying to state a clear goal, what did I want to get back from it.

A challenge I faced was the json passing, I got a little stuck figuring out how the model could read my data in. What I love about the no code solution is you get to see both the input & output at each stage as well as interact with an Ask AI at the same time for any debugging issues.



Connecting Gemini Model was fairly smooth sailing, the details were found [here](#).

I went for the free one of course. There are of course pro's and cons to different models used and limitations which we wont get into for now.



From there we get this big clunky output, structured based on my prompt for a summary, what to focus on etc.

Embarrassingly it says i do not train as much as i should. Well let me tell you coach, I will be running another ultra, training or no training in 2 weeks time.

Then I pass that output and make it into a clear digestible HTML format.  
Maybe I should have the model do this step too? Who knows, least it  
has a bit of consistency then to it.

Finally just needed to set up the email alerting. Of course I didn't know what SMTP details I needed to load, so fortunately the AI assistant could help with that.

You

How do I set up the credentials for SMTP?

**Assistant**

To set up SMTP credentials in n8n, follow these steps:

1. Go to Credentials in n8n and create a new "Send Email" credential.
2. Enter the following details:
  - User: Your email address.
  - Password: Your email password or an app password (some providers require this).
  - Host: The SMTP server address (e.g., smtp.gmail.com for Gmail, smtp-mail.outlook.com for Outlook).
  - Port: Common ports are 465 (SSL), 587 (TLS), or 25 (no encryption). Check your provider's documentation.
  - SSL/TLS: Enable this if your provider requires a secure connection.
  - Disable STARTTLS: Only enable if your provider instructs you to.
  - Client Host Name: Only needed if your provider specifies.
3. Save the credential and use it in your Send Email node.

Provider-specific instructions:

- For Gmail: Enable 2-step verification and generate an app password. Use smtp.gmail.com and port 465 (SSL) or 587 (TLS).
- For Outlook.com: Use your email and password, smtp-mail.outlook.com, and port 587 (SSL/TLS enabled).

If your provider isn't listed, search for their SMTP settings in their documentation.

Did this answer solve your question?

& here is the output!

Strava Summary

to me ➔

Tue, Jun 3 9:36 PM (18 hours ago)

### Ultra-Marathon Training Analysis

**Training Overview**

- Total activities analyzed: 9
- Total mileage: 23.81 km (Runs only)
- Total elevation gain: 21 meters (Runs only)
- Total time (Runs only): 2 hours 34 minutes 56 seconds

**Strengths**

Consistent Pacing: The runner maintains a relatively consistent pace (around 2.5 - 2.75 m/s) across multiple runs, indicating good pacing ability for shorter distances. This demonstrates an ability to manage energy expenditure effectively.

**Areas for Improvement**

- Elevation Training: Very little elevation gain is present in the runs. For ultra-marathon preparation, consistent HR work is essential.
- Workout Variety: The data includes several Workout activities with no distance covered but substantial time and heart rate data, which are not ideal for long distance training. Need for structure endurance sessions.
- Cadence: The runner has a very low average cadence, and should work towards increasing this to reduce strain on the legs and improve efficiency.

It calls me out pretty well for my lack of prep.  
Perhaps I could have done more to set expectations with when the event was, mileage needed per week etc.

#### Race Readiness Assessment

Based on the analyzed data, current preparedness for a 50km ultra event is low. The runner needs to increase weekly mileage, incorporate elevation training, and practice structured long runs to build the necessary endurance and resilience.

#### Periodization Strategy

Given the lack of structured training, a basic periodization plan should start with a base-building phase focusing on increasing weekly mileage gradually, followed by a strength phase including hill workouts, and finally a peaking phase with race-specific long runs.

#### Recommended Next Steps

- Increase Weekly Mileage: Gradually increase total weekly running distance by 10% each week, incorporating longer runs on weekends.
- Incorporate Hill Training: Add one or two weekly workouts focused on hill repeats or sustained climbs.
- Structure Long Runs: Plan long runs with specific pacing targets (e.g., 30 minutes easy pace, 20 minutes tempo pace, followed by remaining distance at easy pace).

On the whole pretty excited to get my first public example going.  
Best part is the workflows are exportable using json and therefore  
shareable with others on N8N.

There are also lots of free templates!

Going further:

Add a trigger event that is based on the next time an activity is logged.

Add in Memory to the agent.

Have a recommended action plan auto push to your calendar.

LOGGING OFF,

CJ

## PLOTTING PENALTIES IN TABLEAU

Hi all,

You may have seen recently **Sara's post** about an easy method for creating penalty shot charts using matplotlib.

```

In [1]: from goal_plot import draw_goal
import pandas as pd
import matplotlib.pyplot as plt

# Import your own shot data here
df = pd.read_csv('my_data.csv')
df['penalty'] = df['penalty'].apply(lambda x: True if x == 'Penalty' else False)
df = df[df['penalty'] == False] # Filter out penalty shots
df = df[['end_location_x', 'end_location_y', 'outcome_name']] # Focus on needed columns
df['class'] = df['penalty'].map({True: 'Penalty', False: 'Shot'}) # Remove shots with missing end location

# Create the figure and axis with dark background
fig, ax = plt.subplots(figsize=(8, 4)) # Create the matplotlib figure and axis
fig.patch.set_facecolor('black') # Set background color for the figure
ax.set_facecolor('black') # Set background color for the plot

# Draw the goal
draw_goal(ax) # Use the draw_goal function from goal_plot

# Plot each penalty shot as a scatter point
for i, row in df[df['penalty']].iterrows(): # Loop through each clean penalty shot
    if row['outcome_name'] == 'Goal': # If row['outcome_name'] == 'Goal' else '#FF4444' # Blue for goal, red for miss
        marker = 'o'
        size = 200
    else:
        marker = 'x'
        size = 200

    # Scatter plot for each shot
    ax.scatter(row['end_location_x'], row['end_location_y'], c='white') # Coordinates of the shot
    marker_color, alpha_val, edgecolor = 'blue', 1.0, 'white' # Style
    if marker == 'x':
        marker_color, alpha_val, edgecolor = 'red', 1.0, 'white' # White border for better visibility

# Save and display the final shot
plt.savefig('goal_draw.png') # Save the figure as an image
plt.show()

```

C:\Users\X\Downloads\Local\Temp\ipykernel\_1220\1220d124899342.ipynb:1: UserWarning: You passed a edgecolor/edgecolors ('white') for an unfilled marker ('o'). Matplotlib is ignoring the edgecolor in favor of the facecolor. This behavior may change in the future.
 ax.scatter(row['end\_location\_x'], row['end\_location\_y'], c='white')

& of course this code is adaptable as you like, but offers a great basis for charting.

You can find a copy of the tutorial and test dataset I created in **my repo.**

1. It imports necessary libraries (goal\_plot for drawing a goal, pandas for data handling, matplotlib for visualization).
2. It loads shot data from a CSV file ('my\_data.csv') and removes any rows with missing location data.
3. It sets up a dark-themed plot (black/dark gray background).
4. It draws a goal outline on the plot using the imported draw\_goal function.
5.
  - Colors goals blue and misses red
  - Uses circle markers (o) for goals and X markers for misses
  - Plots each shot at its end location coordinates (x,y)
6. Finally, it saves the visualization as 'goal\_draw.png' and displays it.

This visualization allows you to see where shots ended up in relation to the goal, with a clear distinction between successful goals and missed shots. The code is using the shot's end\_location\_x and end\_location\_y coordinates for positioning the markers on the plot.

I tried it myself and dumped in a bunch of random points that you can find in the repo, and created the following.



Now if we want to recreate this in Tableau we will want to join to the raw data.

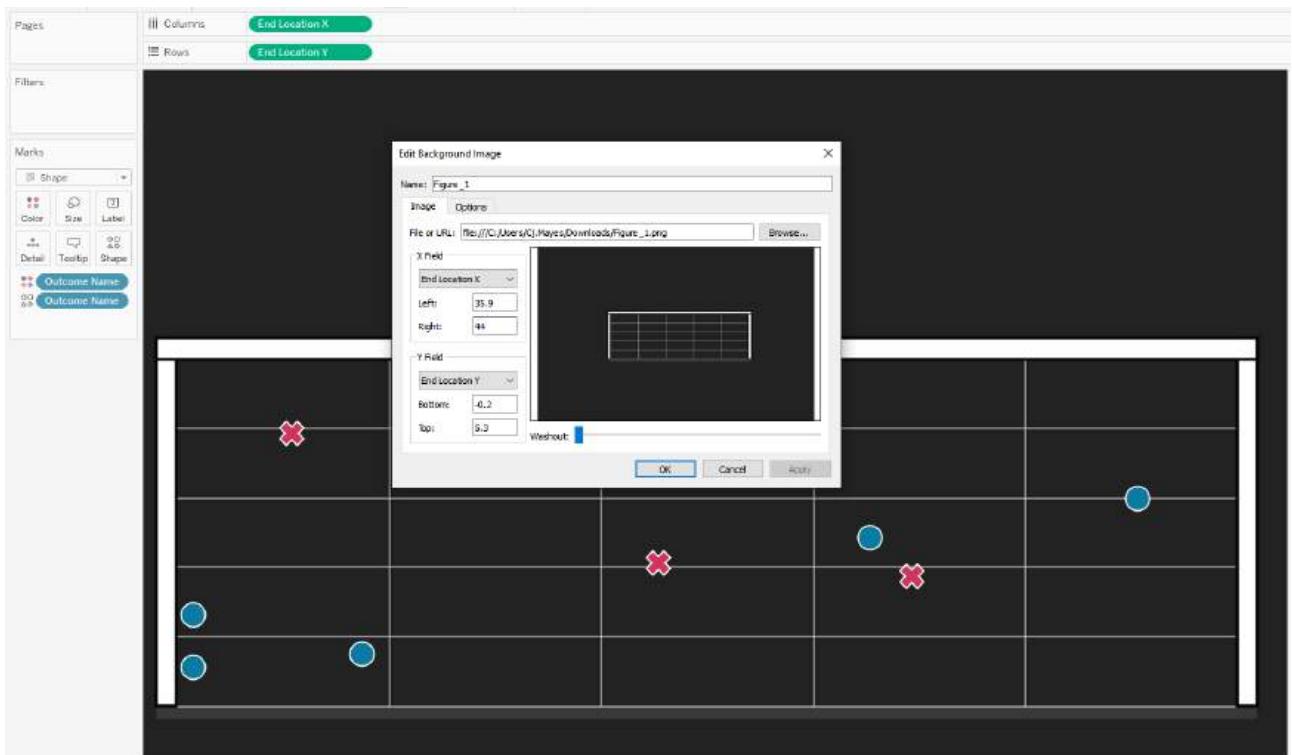
Put End Location X on Columns

Put End Location Y on Rows

Outcome Name on colour and Shape.

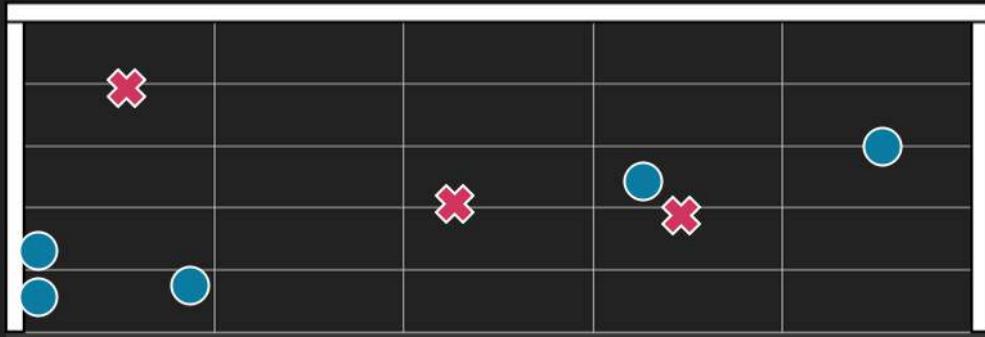
We will want to configure the background map.

Go to Map -> Background images.



Configure the x and y co ordinates to scale to the same values that you have within your dataset.

Nice and simple.



It can get a little bit tricky with the conversions. Firstly, the x,y,z end location scale of your original data, you will want it to fit that of the package.

I saved the goal as an image without any points for when loading it into Tableau as the background image, you can do this just by running the code but excluding the scatter point lines.

Then in background images of Tableau its a little trial and error, until you see that the image you have from plotting it in python matches that on your Tableau screen.

Going further:

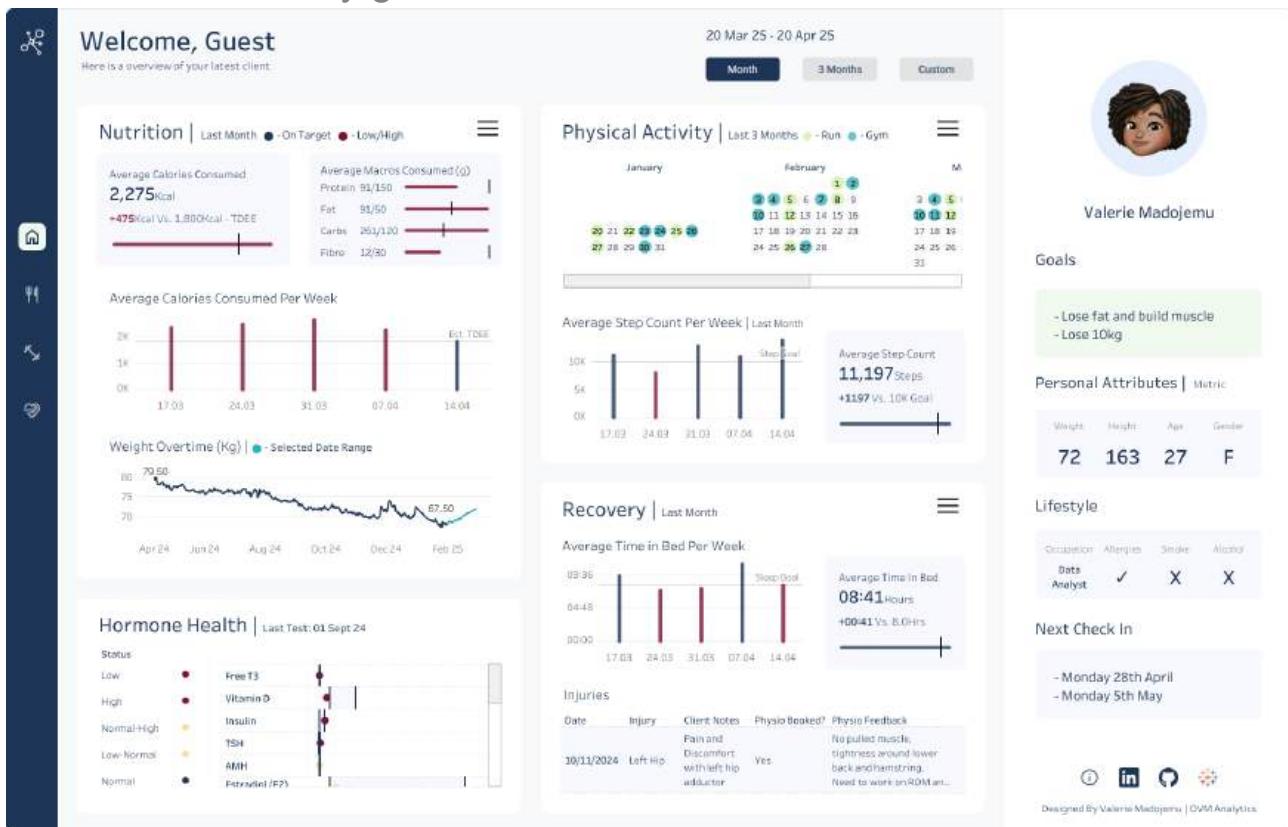
Try use your own data from a recent match.  
Create a different chart for each team in a league.

LOGGING OFF,  
CJ

## HEALTH & FITNESS TRACKER – VALERIE MADOJEMU

Hi all,

This week Valerie Madojemu joins the site. We get to chat about one of her latest dashboard tracking various health metrics. Health, wellness and fitness are most certainly one of my biggest passions so to see Valerie come up with something creative and all encompassing was really great to see. She shares more below.



CJ: Valerie, thank you for giving up time to share more on your portfolio of work. Let's start with a little about you – how did you get into data? Why did you start a public portfolio?



LinkedIn X Twitter

## Valerie Madojemu

Freelance Data Analyst at OVM Analytics LTD. | London, City of, United Kingdom

Follow

Hire Me

Vizzes 31

Favourites 344

Following 45

Followers 140

The screenshot shows Valerie Madojemu's Tableau public portfolio. It features four cards, each displaying a different dashboard:

- Health and Fitness Overview** by Valerie Madojemu: 56 views, 2,877 favorites.
- Hampstead & Westminster Results and Fixtures** by Valerie Madojemu: 21 views, 987 favorites.
- Fleet Information Skytrax Top 10 Airlines 2024** by Valerie Madojemu: 6 views, 260 favorites.
- 2025 Financial Planner** by Valerie Madojemu: 4 views, 254 favorites.

V: I knew I wanted to work in Tech after taking a intro to Python in at university. After I graduated, while working at a coffee shop a colleague recommended I try applying to The Data School (The Information Lab) and third time lucky I got in!

During my time at The Information Lab, I was encouraged to showcase our work through a public portfolio.

At the beginning of my journey, I didn't fully understand the big picture, as I was so focused on learning Tableau, Alteryx, and consulting. But now I understand!

I see it as an opportunity to create opportunities and connect with others in the community.

CJ: Today I want to mainly focus on your Health & Fitness Overview dashboard. We often hear that data becomes valuable when it is combined together to bring insight and action. Can you talk about some of the data sources you used and how others could access the same?

What steps do they need to take?

Yes indeed! Bringing together the data can bring insight! I've worked with nutritionists in the past the focus was on calories, exercise frequency, recovery and more recently looking at hormone health.

## MyFitnessPal (MFP) and Garmin Data

I have been tracking my metrics through MFP for almost 10 Years!

Here I tracked weight and food on a daily basis – which gives my calories, macro and micro nutrients

I have a Garmin watch which tracks my steps and exercises (gym, run, hockey).

My garmin watch can sync with MFP.

MFP has an export feature where I could export my calories, weight and high-level exercise data – three separate tables where the nutrition table is used as a base (as it has highest level of detail in terms of dates, making it easy to relate to the other tables)

## Apple Health and Garmin Data

Finally my sleep data is tracked with sleep app and my Garmin watch, which can be synced to Apple Health and exported to CSV via an App called Auto Export (which offers a week free trial) or you can export directly from Apple Health but you'd have to do additional data prep as the export format is an XML.

1 table was outputted here and a relationship onto the base table, again via date

## Additional data – Manually Created

### Body Measurements

I collected body measurements, hip, waist, bust, and arm every week, however, I felt that for this dashboard, it didn't need to be included in the overview dashboard.

### Injury Data

I manually created a simple table and a relationship to the base table in Tableau

### Hormone Data

I utilised Claude AI to create a fake dataset for my hormone levels and whether it's within range or not.

These three data sets were then related to the base nutrition data via date

CJ: There is also an element of mock data. You mentioned in your linkedin post using Claude AI. How did you go about creating this mock

data through prompts? What did you take into consideration when deciding what biomarkers were most important for the overview page?

The screenshot shows a web-based dashboard titled "Women's Hormone Health Dashboard". A sidebar on the left contains a message from Claude:

I am creating a fitness and nutrition dashboard and want to include some hormone health markers. I would want to present my results and measure it against what the benchmark should be.

What sorts of hormones would be tested for a women and create a fictitious dataset based of 1 women where some hormones would be too low and other in the right range.

The main panel displays a table of hormone data:

Hormone	My_Results_Unit	Benchmark_Min	Benchmark_Max	Status	
Estradiol	(E2),45 pg/mL	30,400	Low-Normal Estrone (E1),10 pg/mL	17,200	Normal Estradiol (E2),8 pg/mL
Total Estrogen	(E3),8 pg/mL	0,10	Normal Total Estrogens,83 pg/mL	61	437,Low-Normal Progesterone
Progesterone	(P),0.2 ng/mL	0,15	Low-Normal Testosterone,35 ng/dL	15,70	Normal TSH,4.8 mU/L
TSH	0.4,4.0	2,3	High Free T3,2.1 pg/mL	2,34,2	Low Free T4,1.1 ng/dL
Cortisol	,0.8,1.8	1,1	Normal High DHEA-S,175 pg/dL	45,380	Normal FSH,7.5 mU/mL
DHEA-S	3.5,12.5	1,2	Normal LH,5.2 mU/mL	2,4,12.6	Normal Insulin,17 μU/mL
Vitamin D	30,100	10	Low Prolactin,12 ng/mL	4,23	Normal AMH,1.2 ng/mL
AMH	1,0,3.5	0,3	Low-Normal		

A note at the bottom says: Let me create a fictional dataset for a 34-year-old woman with some values in normal range.

V: **Here** is the prompt I started with and you can follow the prompting journey.

Prompt:

I am creating a fitness and nutrition dashboard and want to include some hormone health markers, I would want to present my results and measure it against what the benchmark should be.

*What sorts of hormones would be tested for a women and create a fictitious dataset for 1 women where some hormones would be too low and other in the right range*

I don't know what are the most important health markers for a woman, but when I saw the list Claude produced I recognised a few such as

Estrogen, Progesterone, Testosterone and Vitamin D.

For a personal project, I settled with this list, if this was for a client of course I would have done a more thorough check and ask health professionals what are the key hormones.

I put a note in the dashboard that the hormone list was created by AI.

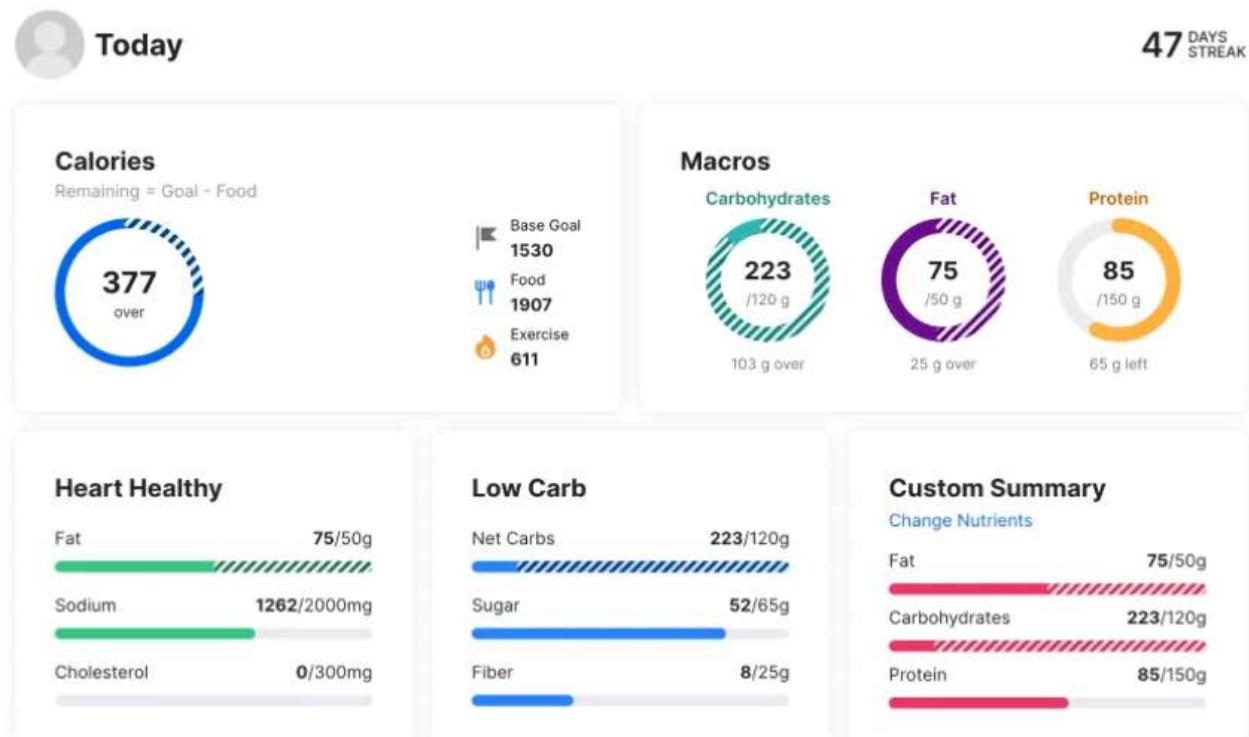
CJ: Looking at wider dashboards, community work, activity tracking apps and health software UI. What impact did those have as inspiration design?

V: Since **Mazeen Muhammed** made his Fitness Dashboard I've wanted to make something similar, an easy to read fitness overview.

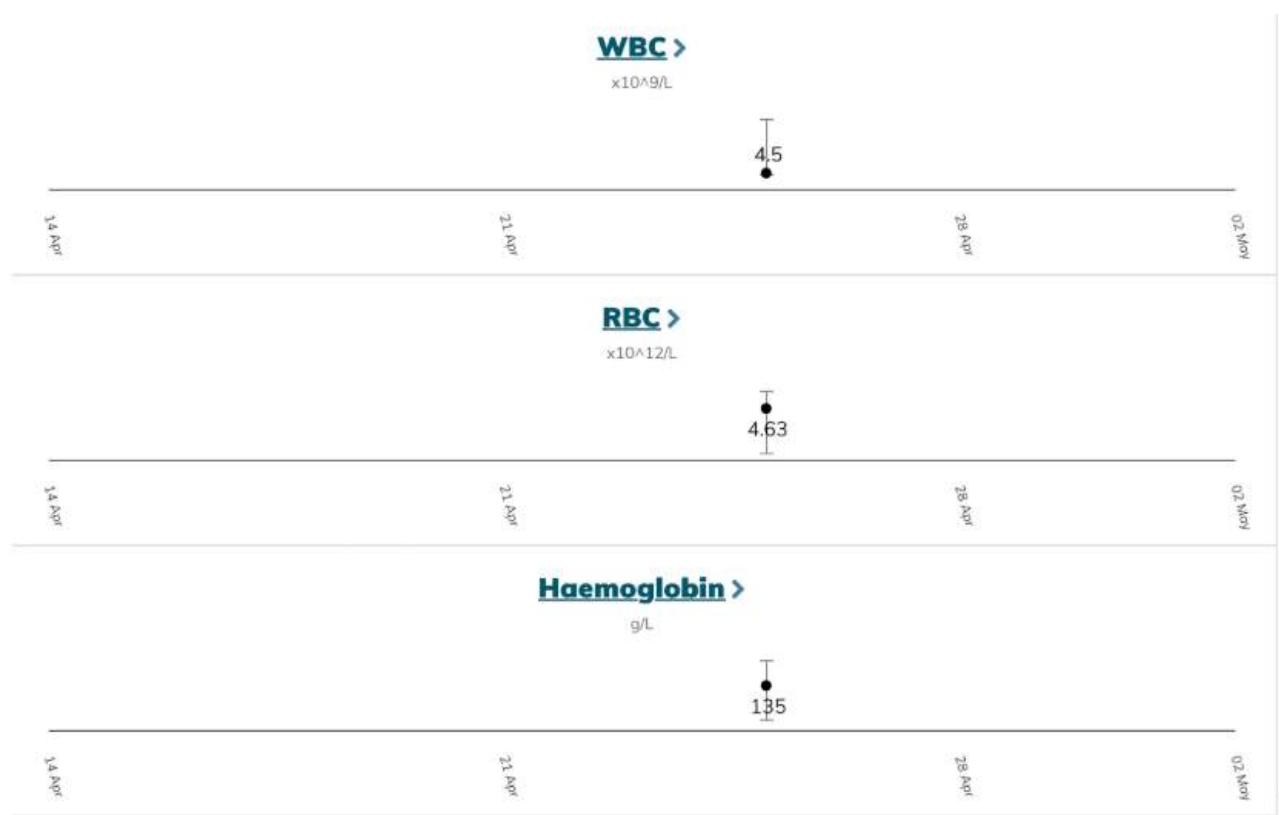


## Charts

For the graphs I took some inspiration from MyFitnessPal, most of my charts are circular bar charts



To plot the hormone data, I actually recently got some blood tests done and liked the box plot format they used and took inspiration there.



## Layout

The layout of the dashboard I took inspiration from these dashboards from Pinterest



CJ: One of your most favoured visuals on Tableau Public is on **UX design Laws**. Could you walk us through how a few of them apply to your Health & Fitness Overview?

# UX Design Laws

[Fitt's Law >](#)

This dashboard highlights essential UX design principles to help data visualisation enthusiasts create user-friendly and impactful visualisations. Each principle is accompanied by practical insights to enhance the design of dashboards and data-driven tools. Apply these principles to craft visualisations that are clear, efficient, and memorable.

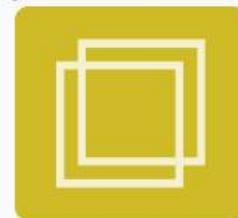
01  
Fitt's Law



02  
Hick's Law



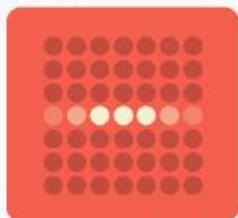
03  
Jakob's Law



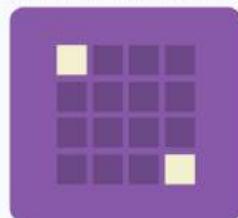
04  
Law of Proximity



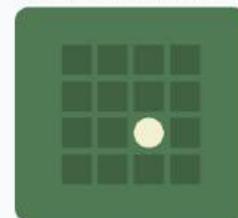
05  
Miller's Law



06  
Serial Position Effect



07  
Von Restorff Effect



08  
Zeigarnik Effect



Designed By Valerie Madjoemnu | @Mad4Viz

Source: lawsofux.

The goal was to make the dashboard as simple as possible so if I were to present this to a PT or nutritionist they could quickly understand what my habits are.

## Hick's Law

I wanted to quickly be able to jump between different time periods such as last 30days, last 90 days and then give the user to decide their own date range, so created buttons to increase the time

## Jakob's Law

I applied Jakob's Law to make sure graphs which are well known and easy to understand, such as bar charts and line charts.

## Law of Proximity

I grouped relevant topics together e.g. associated sleep and injury with Recovery

## Miller's Law

I admit I went over 5-7 visualisation limit. But I chose to break this rule due to the purpose of the dashboard, to look at the overview of each client. Focusing on nutrition, hormone health, physical activity and recovery, to quickly identify any patterns.

From experience I know that if I don't sleep well I overeat, and increase in weight. If weight goes up or down or there's an increase in injuries, I wanted the health professional to find out why it might be happening investigate further.

### Zeigarnik Effect

Health targets on this dashboards have been designed to be binary, for example if I went over the estimated TDEE (Total Daily Energy Expenditure) for a period of time, I would expect weight gain to occur.

The bullet charts help illustrate by how much and colour to highlight.

CJ: You've added elements of flexibility in changing targets into your dashboard. Why is that important compared to just tracking metrics over time. Were there any insights to some of your "real" data that you were surprised at?

V: Habits and goals change over time. For example, if weight loss was the goal, the client wants to be in a calorie deficit, hitting protein goals and hitting the gym/exercise consistently.

If weight loss doesn't occur, I wanted to be able to quickly troubleshoot the issue by investigating if the client overate; or re-affirm the client with reassuring that they have built muscle and need to be patient in their weight loss journey.

Alternatively if a client just finished a football season and wanted to maintain their physique or work on something particular in the off-season, then the targets can be set to maintain the athletes physique on target for pre-season (and include breaks) or make sure they are fuelled enough to accomplish their goals.

I guess I would say I was surprised at how "low" my maintenance calories was. I think it is easy to get influenced on social media thinking I needed more than this when the reality is most of my hours are sat at a Laptop most of the day. I've quickly become obsessed with making sure I get in my steps.

CJ: How has the **UI design principles** you've referenced in past work helped elevate your work in recent vizzes and particularly your health and fitness visual?

# UI Design Principles

[Next Page >](#)

This dashboard highlights essential UI design principles to help data enthusiasts create visually appealing and impactful visualisations. Each principle enhances the **Aesthetic Usability Effect** - where users perceive attractive designs as more usable. When designs create positive emotional responses, users believe they work better. Apply these principles to craft visualisations that increase engagement and improve data comprehension.

01

Visual Hierarchy  
& Layout

02

Typography  
& Scale

03

Colour Theory  
& Contrast

04

White Space  
& Balance



05

Consistency  
& Design

06

Iconography  
& Design

07

Alignment  
& Proximity

08

Feedback  
& Affordance

Designed By Valerie Madojemu | @Mad4Viz [in](#) [X](#) [o](#)

V: These design principles have influenced me a lot, making sure items on a dashboard are easy to interpret and have a clean aesthetic feel.

## Visual Hierarchy

The layout was influenced by following the Z-Pattern, so you would view nutrition first and recovery last

## Colour Theory

Picking a colour palette is my weakness, but what helped me was watching **Kimly Scott TC25** talk where she talked about creating a simple monochromatic colour scheme for one of her visualisations. I know my colour scheme is not monochromatic however, I didn't want to use a super high contrast colour, originally I thought I'd use orange to highlight then a target has not been hit, but instead opted for a dark red.

## Whitespace & Alignment and Proximity

With a loaded dashboard, I wanted to make sure each section had enough whitespace between them for the user to digest each section.

## Iconography

With the legends I made sure to keep the colour dots circles across the dashboard.

## Feedback & Affordance

Different parts of the dashboard are controlled by different date ranges.

I made it clear through various ways:

- Date Buttons – turning blue when selected and creating a MIN and MAX date for each date range selected
- Date range within title
- Use of axis on graphs

CJ Round-up:

Thanks for sharing some of your design tips in relation to your work.

Check out Valeries post [here](#).

“Your habits are the silent architects of your life.”

LOGGING OFF,

CJ

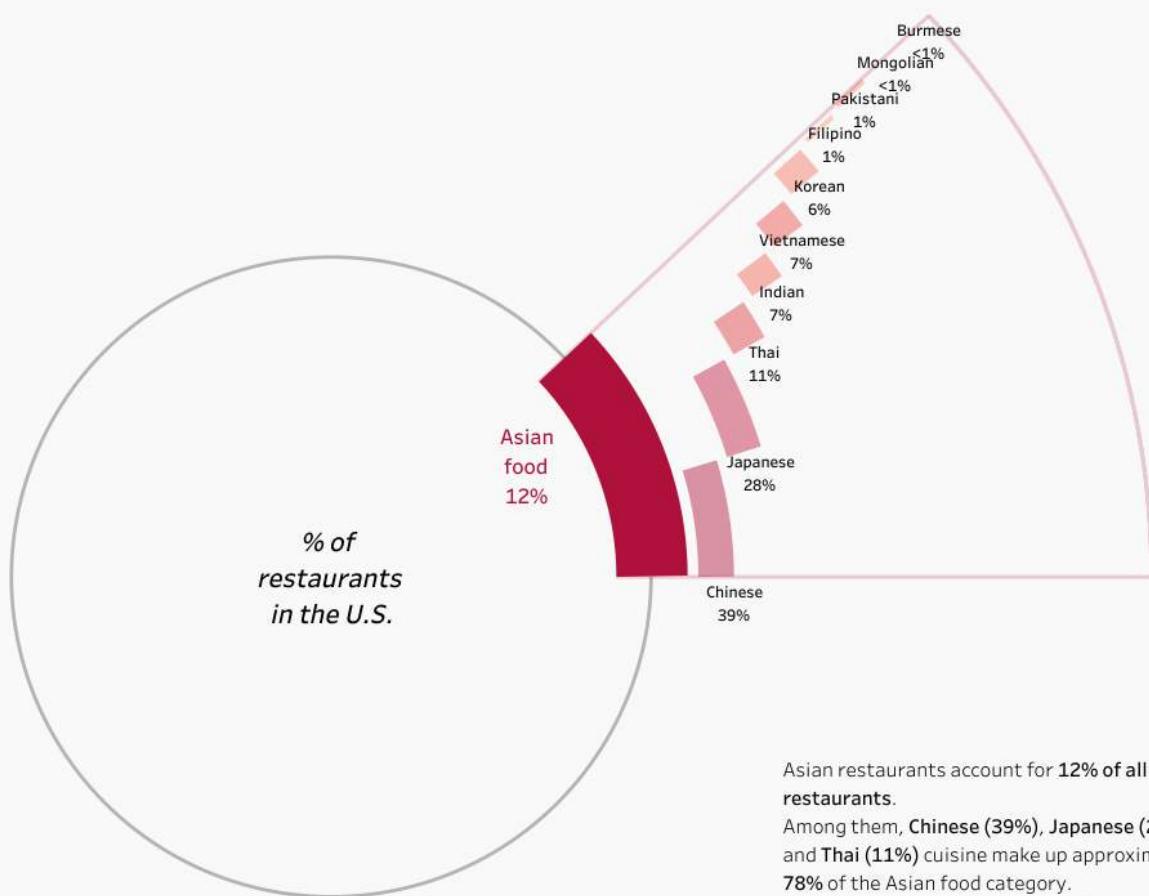
## RADIAL WATERFALL CHART – KYEONGSEOK MIN

Hi all,

Something creatively bespoke to share today, coined a radial waterfall chart. I love when i see people build pretty adventurous chart types and are willing to shift away from the bars and line charts that we all know are best using in the business setting.

# Inside the 12%: The Landscape of Asian Restaurants

A radial snapshot of Asian restaurant diversity



Makeover Monday 2025 Week 19

Challenge by Kyeongseok MIN (@WXMINTAB)

I'm really pleased that Kyeongseok MIN joins the blog today.  
You can find **Kyeongseok MIN on X, linkedin and tableau public.**

After completing his bachelor's degree, Kyeongseok Min started his career as a research associate at the National Institute of Meteorological Sciences. There, he focused on data processing and visualization, simulation analysis, and supporting impact-based forecasts. After two years, Kyeongseok Min pursued and obtained a PhD in Science, specializing in meteorology, with a focus on linear precipitation systems and severe rainfall. Upon joining his current company, Kyeongseok Min transitioned to a data scientist role, contributing to digital transformation

(DX) through data analysis and Tableau visualizations to support client decision-making.

Over to you!

1.

#### Dataset Origin

This visualization was created for **#MakeoverMonday 2025 Week 19** challenge, using a dataset that shows the percentage of Asian restaurants among all U.S. restaurants.

The original data can be found [here](#).

	A	B	C
1	Category	% of Asian Restaurants	
2	Chinese	39	
3	Japanese	28	
4	Thai	11	
5	Indian	7	
6	Vietnamese	7	
7	Korean	6	
8	Filipino	1	
9	Pakistani	1	
10	Mongolian	<1	
11	Burmese	<1	
12	Other Asian/Un:	19	
13			
14			

Supplemental data: Manually generated sequence from 1 to 401 for radial bar pathing

	A
1	1
2	2
3	3
4	4
5	5
6	6
7	7
8	8
9	9
10	10
11	11
12	12
13	13
14	14
15	15
16	16

2.

### **Why this chart type?**

I chose a donut chart to highlight one key point:  
Asian restaurants make up 12% of all U.S. restaurants.  
To drill down into that 12%, I used a radial waterfall-style chart to show  
how much each category contributes.  
I thought this would be a fun and unique way to visualize the variety of  
Asian restaurants.

3.

### **Data Preparation**

The original dataset had two sheets:  
1)The share of Asian restaurants among all U.S. restaurants (Category -> Total)  
2)To drill down into Asian restaurant types



Asian — Total\_rest.

How do relationships differ from joins? Learn more

Asian      Operator      Total\_rest.

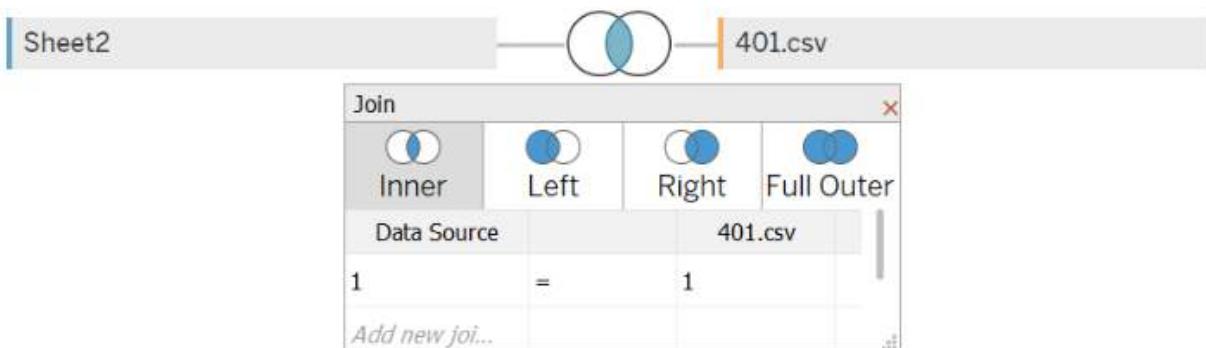
"Asian food (total)"    <--    **401.csv**

Add more fields

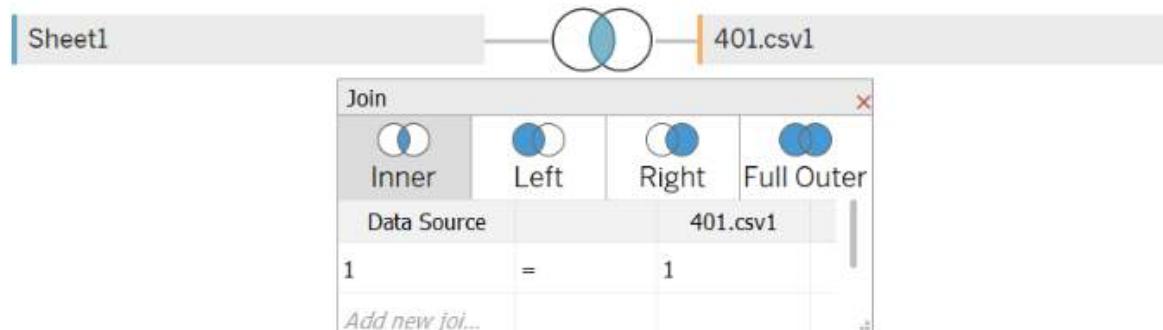
Alt Sheet2	Alt Sheet2	# Calculation	#	Calculation	#
Category	% of Asian Restaurants	Asian_Rest_Rank	AIRLctv	Calculation	Label_pos
Chinese	39	1	1	1.13000	1
Japanese	28	2	1	1.26000	1
Thai	11	3	1	1.39000	1
Indian	7	4	1	1.52000	1

I joined these two sheets in Tableau by matching the “Asian food (total)” category from sheet 1 and creating a calculated field with the same label in sheet 2.

Asian is made of 2 tables. ⓘ



Total\_rest. is made of 2 tables. ⓘ

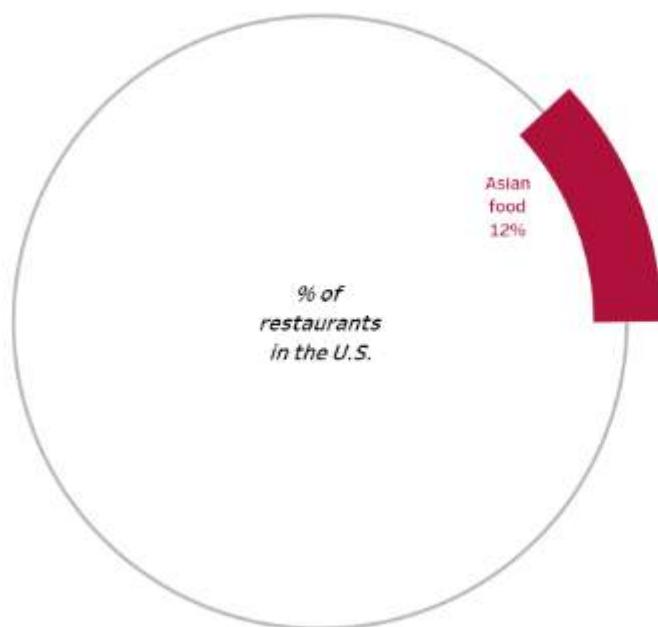


To draw the donut and curved bars, I also used a helper table (401.csv) containing values from 1 to 401. This table was an inner join with both sheets.

4.

### How the Chart Was Built

#### 4.1 Drawing the Outer Circle



To create the donut shape, I used two MAKEPOINT calculations:

A thin ring to display the full circle

A thicker segment to highlight the 12% portion occupied by Asian restaurants

Both are based on polar coordinates, using the following formula:

MAKEPOINT(diameter \* SIN(degree), diameter \* COS(degree))

Step 1: Thin Full Circle

I created 401 points using a helper table ( ranging from 1 to 401) and defined the angle and radius as follows:

```
2 * PI() * (
IF <= 200 THEN
    / 200
ELSE
    (200 - ( - 201)) / 200
END
)
```

```

IF <= 200 THEN
  0.9025
ELSE
  0.8925
END

```

MAKEPOINT( \* SIN(), \* COS())

After placing on the Detail shelf:

- 1)Change the mark type to Polygon
- 2)Drag to the Path shelf (as a dimension)

This creates a smooth, closed ring by connecting all the points sequentially.

Step 2: Thick Highlight for Asian Restaurants (12%)

To emphasize the 12% share of Asian restaurants, I adjusted the degree calculation using the percentage value:

```

2*pi()*(

IF <= 200 THEN
  /200
ELSE
  (200 - (-201))/200
END)

*((if = "Asian food (total)" THEN END)/100)

```

MAKEPOINT( \* SIN(), \* COS())

Add a Marks Layer

Polygon

Color: Orange, Size: 10

Path

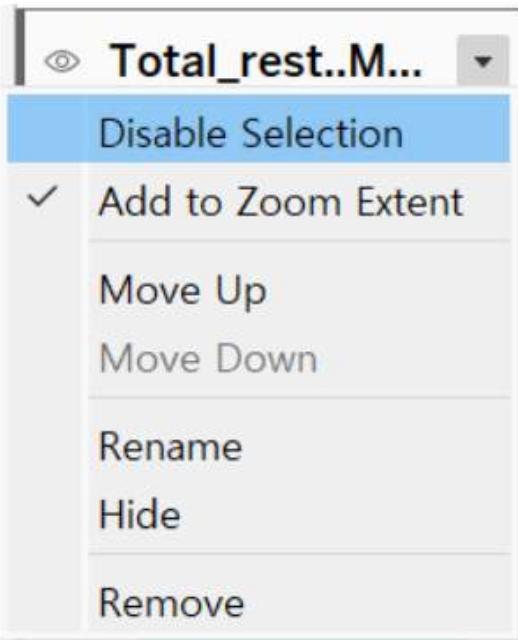
Detail: COLLECT(MP\_TOTAL)

Path: T\_value\_for\_po...

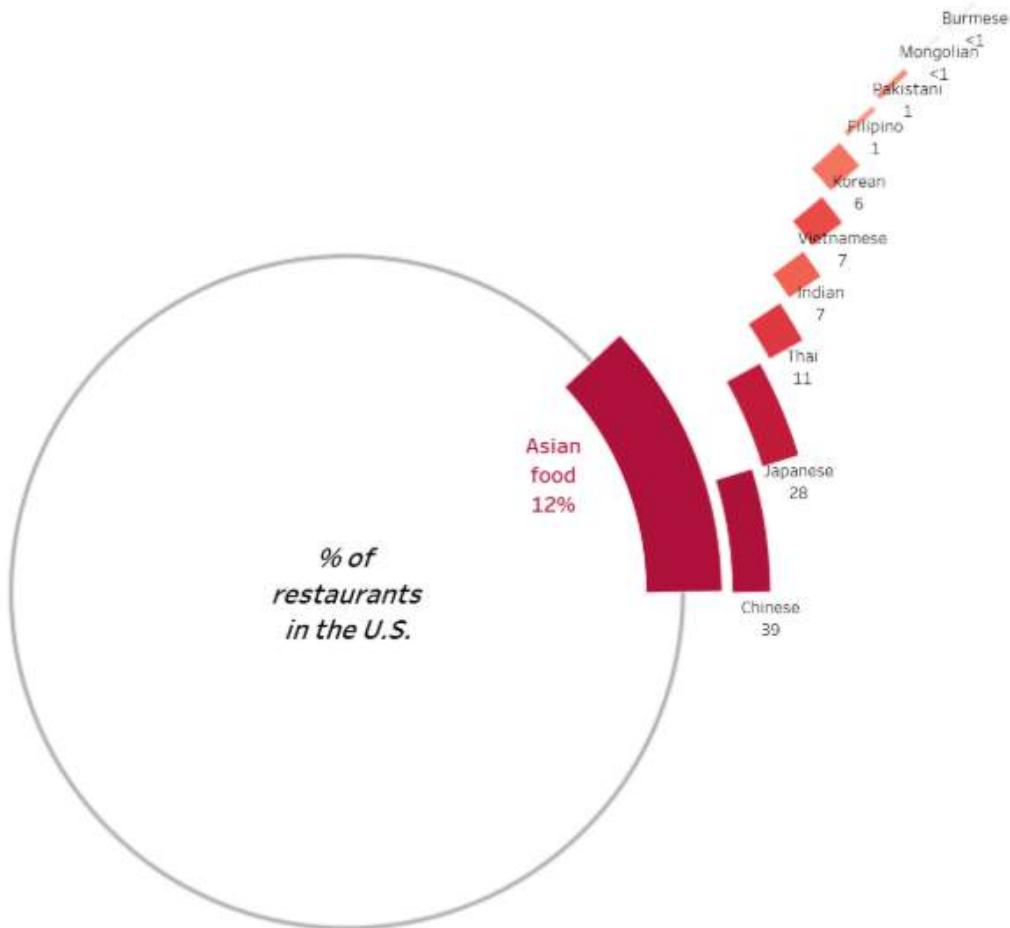
This results in a bold, curved segment that visually emphasizes the presence of Asian restaurants within the overall restaurant landscape.

To add category labels (e.g., “Asian Food”), you can also use MAKEPOINT to position text precisely on the map.

Additionally, if you want to disable selection, check “Disable Selection” for each layer.



#### 4.2 Building the Radial Waterfall (Linear Bar Chart)



To drill down the 12% segment into categories of Asian restaurants, I built a radial waterfall-style chart using MAKEPOINT. This required calculating the accumulated share (running sum) for each category to correctly place the bars in order.

#### Step 1: Assign Category Rank & Calculate the Accumulated Share (Running Sum)

Since MAKEPOINT calculations don't support table calculations like INDEX() or RUNNING\_SUM() directly, I manually assigned each restaurant category a rank (similar to an index) and used that to compute the running sum step by step. This cumulative value was then used to calculate the correct angle for each bar segment in the radial layout.

CASE

WHEN "Chinese" THEN 1

```
WHEN "Japanese" THEN 2
    WHEN "Thai" THEN 3
    WHEN "Indian" THEN 4
WHEN "Vietnamese" THEN 5
    WHEN "Korean" THEN 6
    WHEN "Filipino" THEN 7
    WHEN "Pakistani" THEN 8
    WHEN "Mongolian" THEN 9
    WHEN "Burmese" THEN 10
        END
```

okay and next quite a chunky calculation.....

```
CASE
    WHEN 1 THEN 0
    WHEN 2 THEN {MAX(IF =1 THEN END)}
    WHEN 3 THEN {MAX(IF =1 THEN END)}+
        {MAX(IF =2 THEN END)}
    WHEN 4 THEN {MAX(IF =1 THEN END)}+
        {MAX(IF =2 THEN END)}+
        {MAX(IF =3 THEN END)}
    WHEN 5 THEN {MAX(IF =1 THEN END)}+
        {MAX(IF =2 THEN END)}+
        {MAX(IF =3 THEN END)}+
        {MAX(IF =4 THEN END)}
    WHEN 6 THEN {MAX(IF =1 THEN END)}+
        {MAX(IF =2 THEN END)}+
        {MAX(IF =3 THEN END)}+
        {MAX(IF =4 THEN END)}+
        {MAX(IF =5 THEN END)}
    WHEN 7 THEN {MAX(IF =1 THEN END)}+
        {MAX(IF =2 THEN END)}+
        {MAX(IF =3 THEN END)}+
        {MAX(IF =4 THEN END)}+
        {MAX(IF =5 THEN END)}+
        {MAX(IF =6 THEN END)}
```

```

WHEN 8 THEN {MAX(IF =1 THEN END)}+
{MAX(IF =2 THEN END)}+
{MAX(IF =3 THEN END)}+
{MAX(IF =4 THEN END)}+
{MAX(IF =5 THEN END)}+
{MAX(IF =6 THEN END)}+
{MAX(IF =7 THEN END)}

WHEN 9 THEN {MAX(IF =1 THEN END)}+
{MAX(IF =2 THEN END)}+
{MAX(IF =3 THEN END)}+
{MAX(IF =4 THEN END)}+
{MAX(IF =5 THEN END)}+
{MAX(IF =6 THEN END)}+
{MAX(IF =7 THEN END)}+
{MAX(IF =8 THEN END)}

WHEN 10 THEN {MAX(IF =1 THEN END)}+
{MAX(IF =2 THEN END)}+
{MAX(IF =3 THEN END)}+
{MAX(IF =4 THEN END)}+
{MAX(IF =5 THEN END)}+
{MAX(IF =6 THEN END)}+
{MAX(IF =7 THEN END)}+
{MAX(IF =8 THEN END)}+
{MAX(IF =9 THEN END)}

END

```

### Step 2: Define Total Angle Span (12%)

$$2 * \text{PI()} * ( / 100)$$

The field defines the total angle span representing the full 12% share of Asian restaurants.

Although this value isn't explicitly filtered in the calculation, it works as intended because in the data preparation step, the row for "Asian food (total)" was already joined with each category. This ensures that the value is automatically associated and used correctly during the angle calculation.

### Step 3: Compute Angles and Radius per Category

Each bar's angle and its base position were calculated using a helper table ( from 1 to 401) and the running total.

Angle within segment (bar width):

```
* (
  IF <= 200 THEN
    / 200
  ELSE
    (200 - (- 201)) / 200
  END
)*
```

Starting angle for each bar (cumulative base):

```
* (
  IF <= 200 THEN
    / 200
  ELSE
    (200 - (- 201)) / 200
  END
)*
```

Final angle position:

+

Radius increases by category rank (spread out effect):

```
IF <= 200 THEN
  1 + * ( / 10)
ELSE
  1 + * ( / 10) - 0.1
END
```

is a float parameter (I set it to 1.3).

### Step 4: Create MAKEPOINT for Bars

Use the values above to generate the points for polygon shapes:

MAKEPOINT( \* SIN(), \* COS())

- 1)Add a layer using
- 2)Set mark type to Polygon
- 3)Place on Path

4)Add to Color

#### Step 5: Label Below Each Bar

For clean labeling, place it at the start of each bar

MAKEPOINT( \* SIN(), \* COS())

1)Add a new layer using

2)Mark type: Line

3)Add and to Label

Use this to limit label to final position only:

IF <= 200 THEN ELSE 0 END

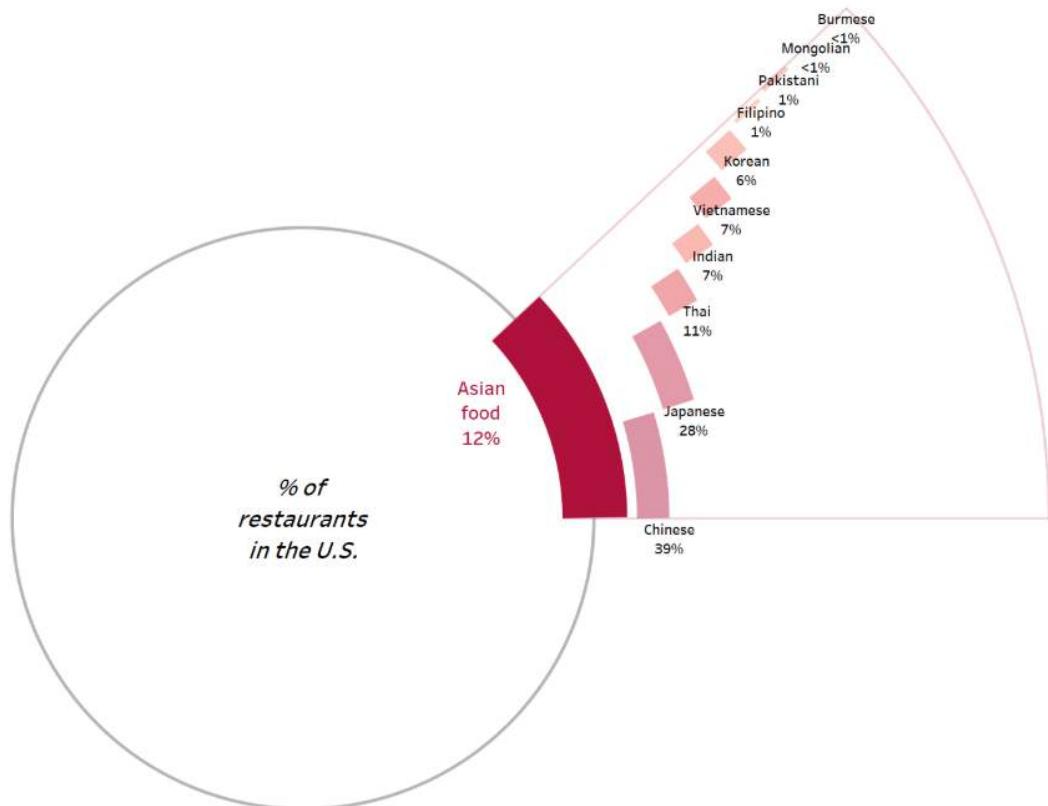
→ Drag to Path

This technique ensures the label appears only once per category at the outer end of the bar, looking like a clean label below the curved bar.

#### Step 6: Final Touches

1)Adjust line color and opacity to hide connector lines if needed

2)Style the polygon colors as preferred for clarity or aesthetics



To complete the donut's appearance, I added outlines using both  
MAKELINE and MAKEPOINT.

## Step 1: Draw Inner and Outer Lines

I created two straight lines to connect the center to the inner and outer radius of the donut segment

```
MAKELINE(MAKEPOINT(0, (1 - 0.2)), MAKEPOINT(0, (1 + )))  
        MAKELINE(  
            MAKEPOINT(0.8 * SIN(), (1 - 0.2) * COS()),  
            MAKEPOINT((1 + ) * SIN(), (1 + ) * COS())  
        )
```

Set the mark type to Map.

Use a similar color as your donut segment to make the outline blend smoothly.

## Step 2: Draw the Outer Curve

This is a curved line that outlines the donut's outer edge:

```
MAKEPOINT((1 + ) * SIN(), (1 + ) * COS())
```

1)Add this as a new layer.

2)Set the mark type to Line.

3)Add to the Path shelf to connect the points smoothly.

4)Match the color to your donut chart to keep visual consistency.

## 4.4 (Optional) Interactive Toggle Using Parameters

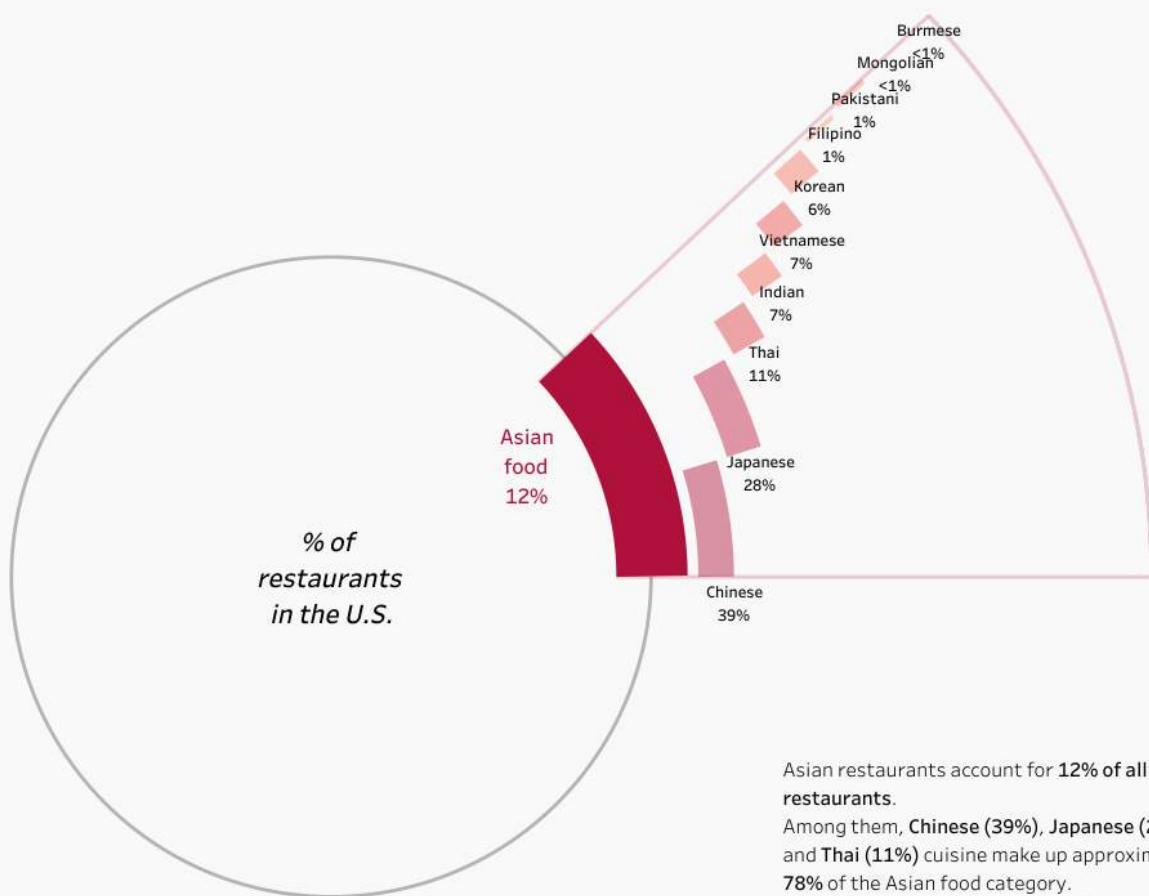
To make the chart expandable and collapsible, I used a parameter-based toggle. Curious how it works? Download my dashboard (.twbx) on Tableau Public and take a closer look under the hood!

## 4.5 Final Touches

After building all the components, just insert them into the dashboard and customize the layout and design — and you're done!

# Inside the 12%: The Landscape of Asian Restaurants

A radial snapshot of Asian restaurant diversity



Asian restaurants account for **12%** of all U.S. restaurants.

Among them, Chinese (39%), Japanese (28%), and Thai (11%) cuisine make up approximately **78%** of the Asian food category.

This shows a strong dominance of East and Southeast Asian cuisines in the U.S. food landscape.

Makeover Monday 2025 Week 19

Challenge by Kyeongseok MIN (@WXMINTAB)

I call this style a Radial Waterfall Chart.

CJ Round-up:

A lot of layers, a lot of trigonometry, and a whole lot of fun. Perhaps one for those that want to utilise some of the calculations for their own radial visuals, but I love that whilst this may differ from your standard pie & bar, it still conveyed the same message as part to whole and then the make up of that smaller segment. Almost like a sunburst variation.

LOGGING OFF,

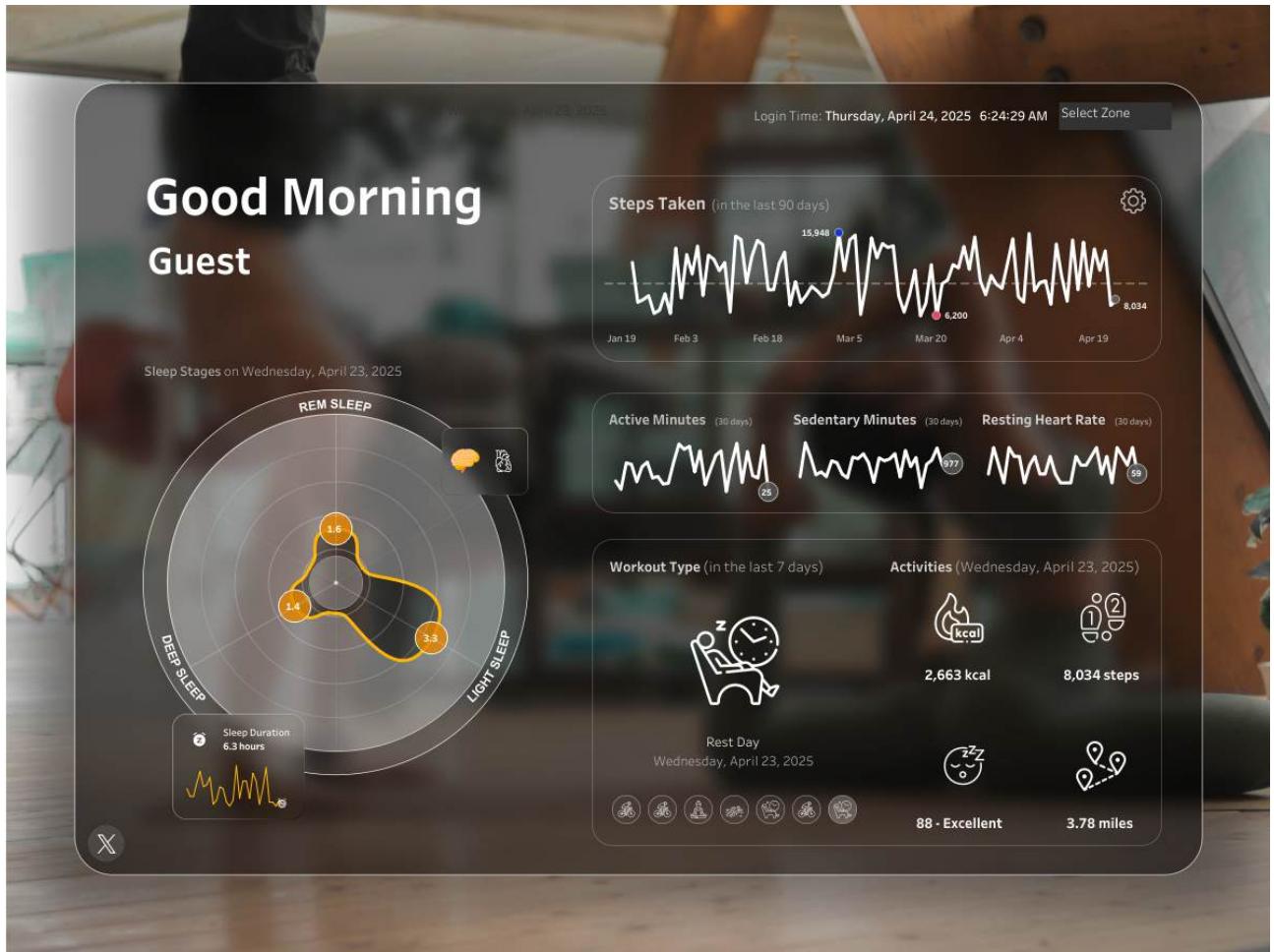
CJ

## INVERSE CURVY RADARS IN TABLEAU – SEUN ADEYEMO

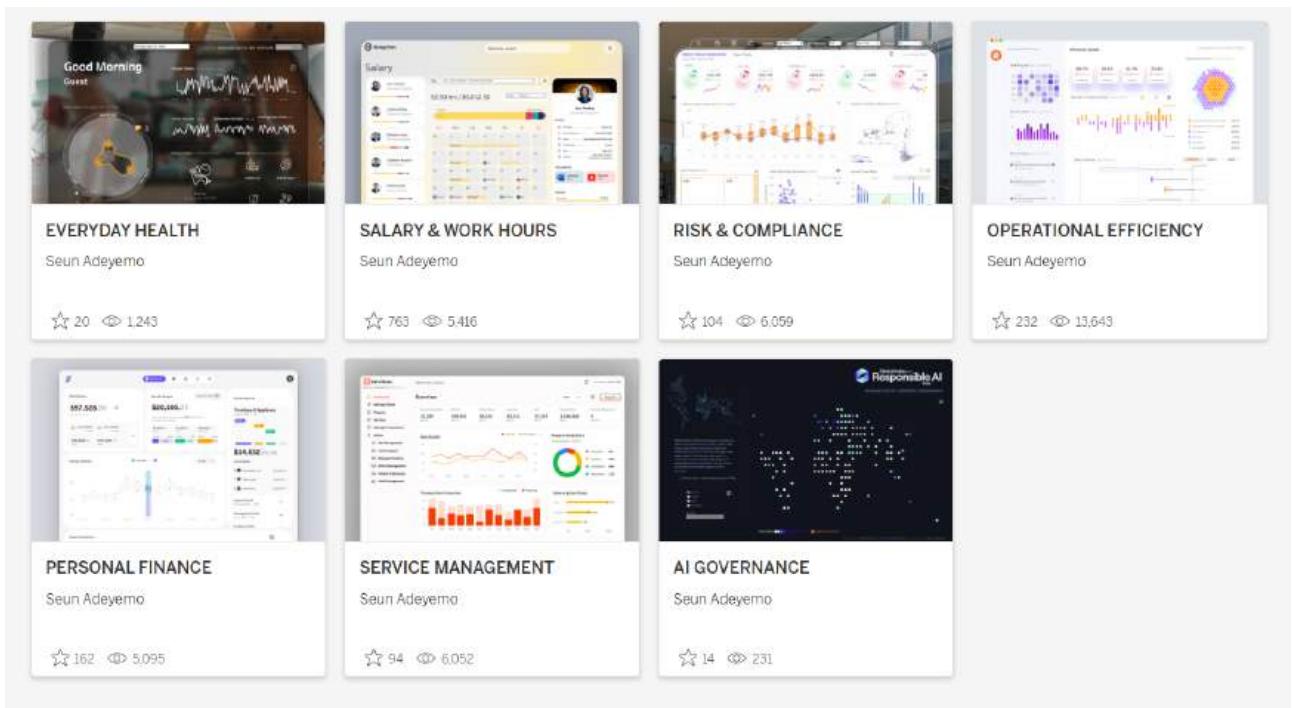
Hi all,

A warm welcome back to another blog, and this week, Seun has kindly offered up his time to share a really neat inverse polygon trick in Tableau.

Just check out this radar, it works so well with the glassmorphism background.



Seuns profile is so visually striking, it is of no surprise he picked up an award at this years Tableau conference. So i'm glad he is here to share his expertise through a **inverse radar template** today.



Before we jump in Seun, for those wanting to learn more about you can you describe how you got into data?

S: Hi! My name is Seun Adeyemo. I'm based in Lagos, Nigeria. I love Art, in fact, it was my first love. As a teenager my passion was drawing (with pencils), painting with water colours (sometimes oil paint), and I did lots of beautiful crafts. After high school, I taught kids with similar passion how to make beautiful stuffs like card greetings, straw mats, cardboard designs, tie and dye, etc.

But i also love Mathematics, because to me it's about solving problems by employing both logical and analytical reasoning. I went on to studying further mathematics, In the end, I produced a **further mathematics textbook!**

After studying Economics at university, I grew more of a passion for graphs, and by chance I found Tableau in 2018. It was the perfect software that allowed me to blend everything together (Art, Math & Economics).

After some practice, I started training university students every year. And you know, you become an expert more by teaching, so here I am now!

Now let's get into the tutorial.

So, how did I develop the first inverse chart of that curvy radar? Sincerely I think the credit needs to go to Ken Flerlage who with Lisa

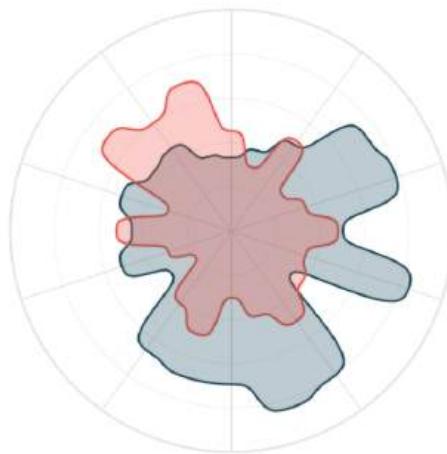
Trescott and Kevin Flerlage created the first curvy radar on Tableau. I was trying to build a curve style radar and so I remembered that I read Ken's article on that sometime ago. So I went back to the article, downloaded the chart and then got to work.

## You can read it here.

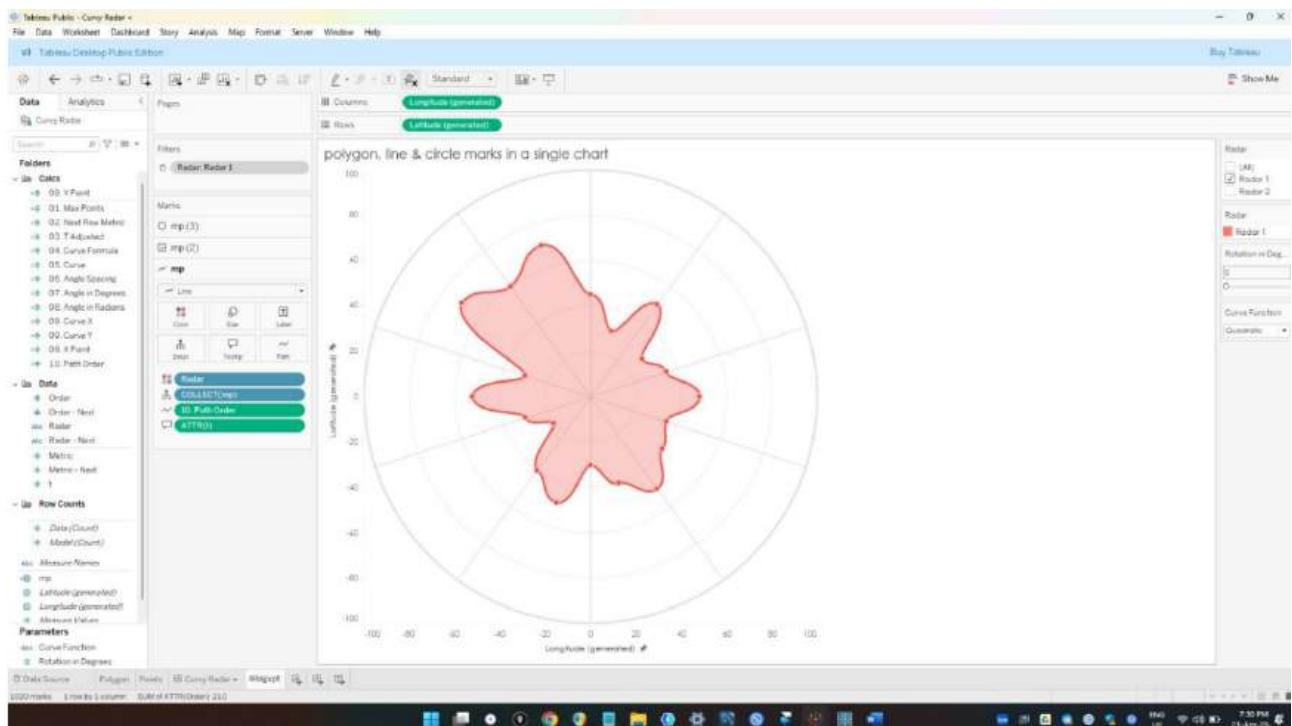
### Automating It

As Lisa mentioned, she leveraged Alexander Varmakov's [brilliant technique](#) for creating a curvy radial bump chart. Of course, she had to modify the technique to handle a radar chart instead of a bump chart, but the concepts are very similar. The bulk of that work was done in Excel. But after a bit of back and forth with Lisa, we wondered if we could push most of this to Tableau. By doing that, we could turn this into an easy-to-use template. So, using some [data densification techniques](#), we were able to transition from the Excel-based solution to one almost entirely done in Tableau. I'll admit that some of Alexander's calculations were a little over my head, so I drew from my past experience of drawing curvy and circular stuff and reinvented a few of the calculations along the way. I also made adjustments to ensure that the last point would connect to the first point, addressing the challenge Lisa noted earlier.

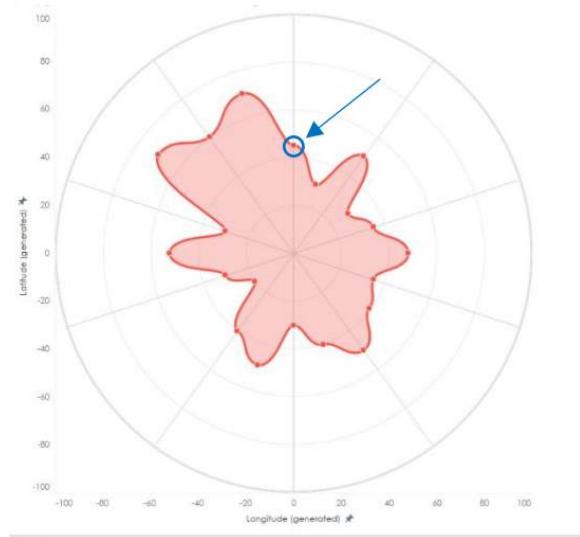
In the end, we were thrilled with the result:



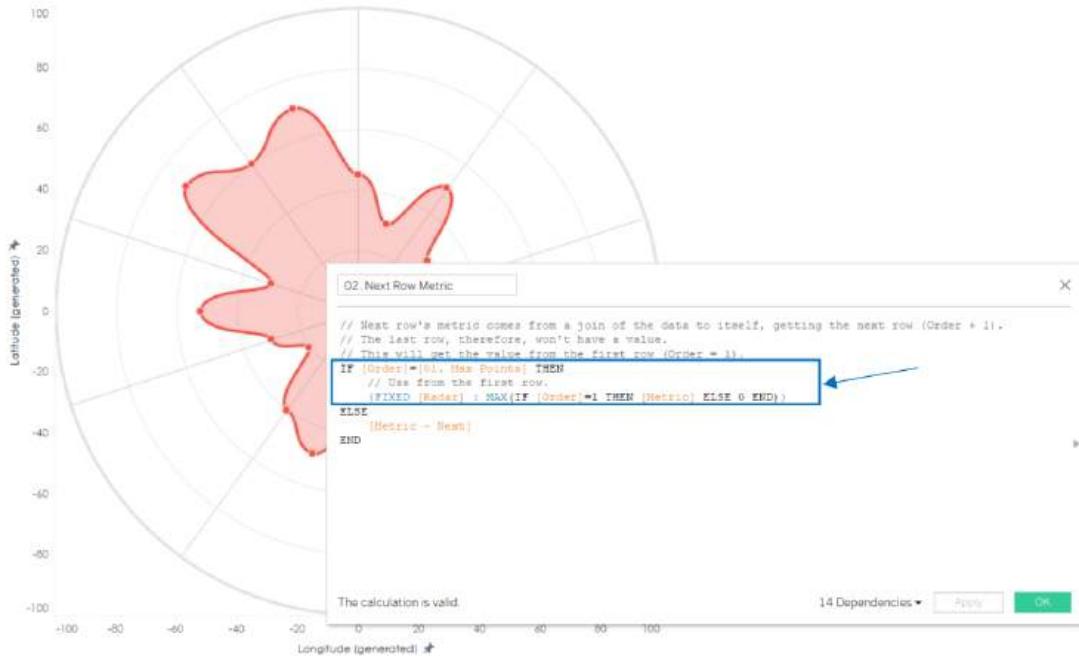
I would save the explanation on how to build a curvy radar as that has already been explained best by Ken in his website. So, I'd just go straight to explaining how I built the inverse of it and created a see-through effect.



So, here's what we got, the Radar 1 from Ken's template. We all know this is a polygon, which means there are points and these points "connect" together, essentially the last data point is connected to the first data point.

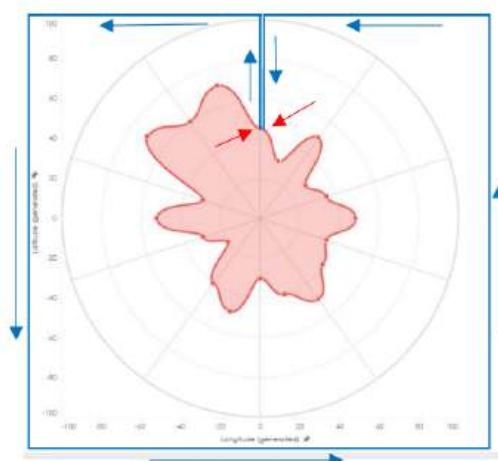


You say, how do I know it's that point? Well Ken's calculated field 02. Next Row Metric says that here!

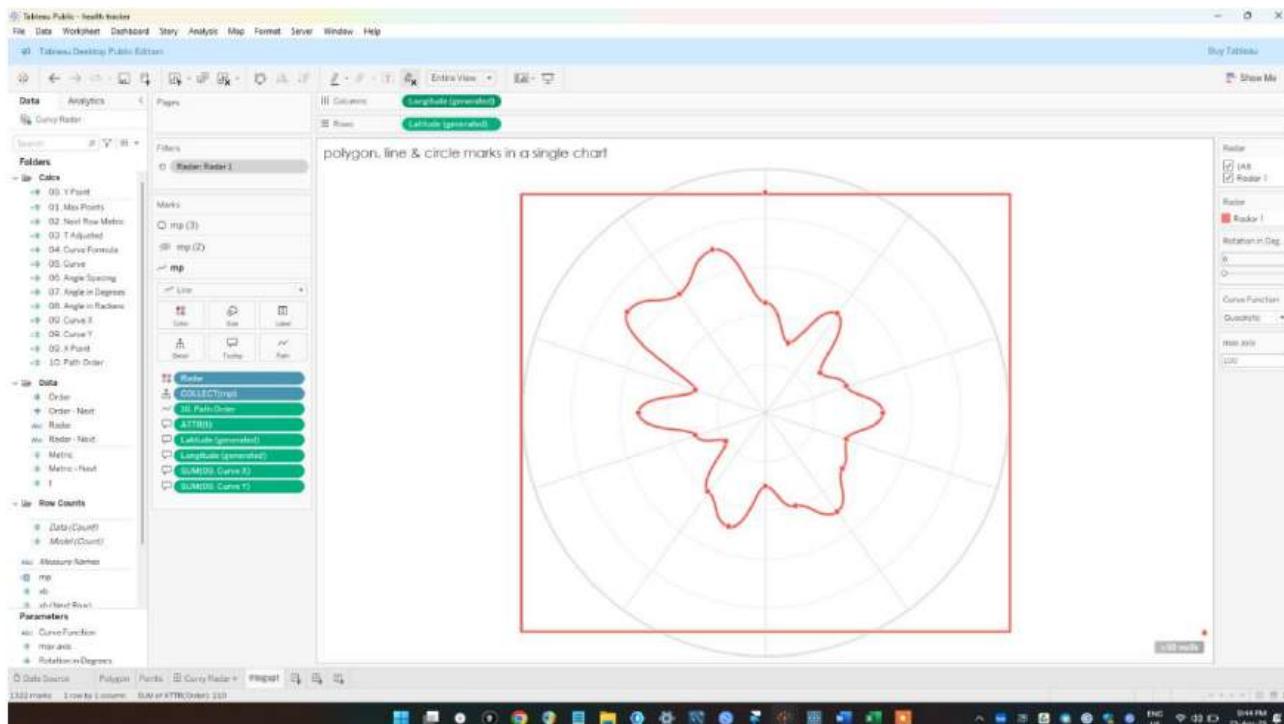


It says if the data point is at its end (max point), the next row should be the value of the first data point (...if order = 1 then metric...), also note this was fixed.

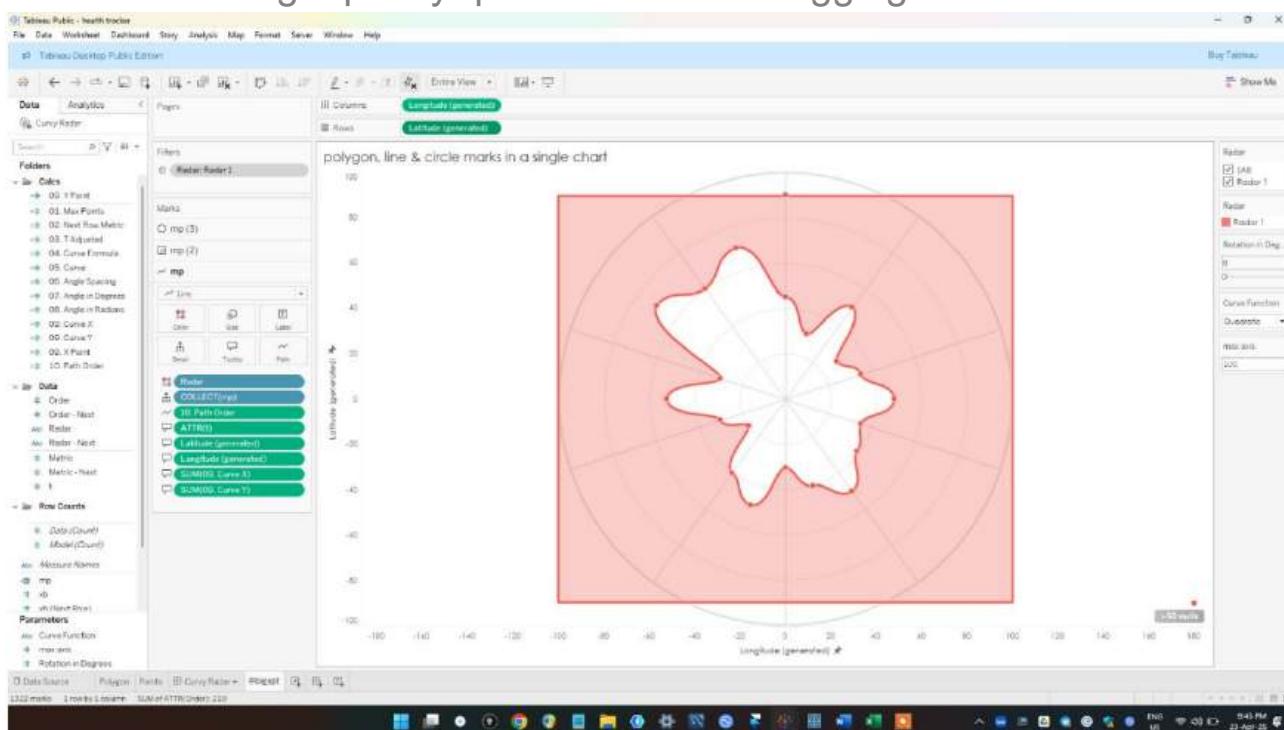
So, I thought, what if we extend this data point from the last point, so that it goes rectangular around the data view and back into the radar shape meeting with the first data point. This was what I had in mind.



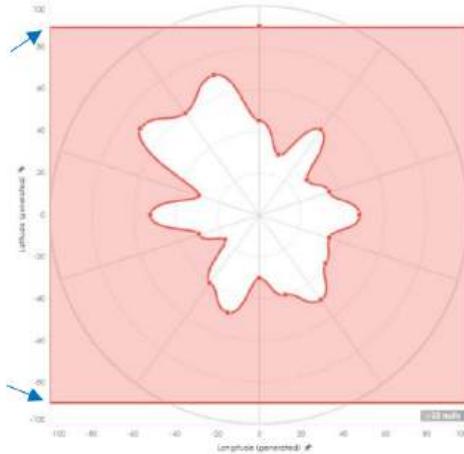
And this was what I eventually built



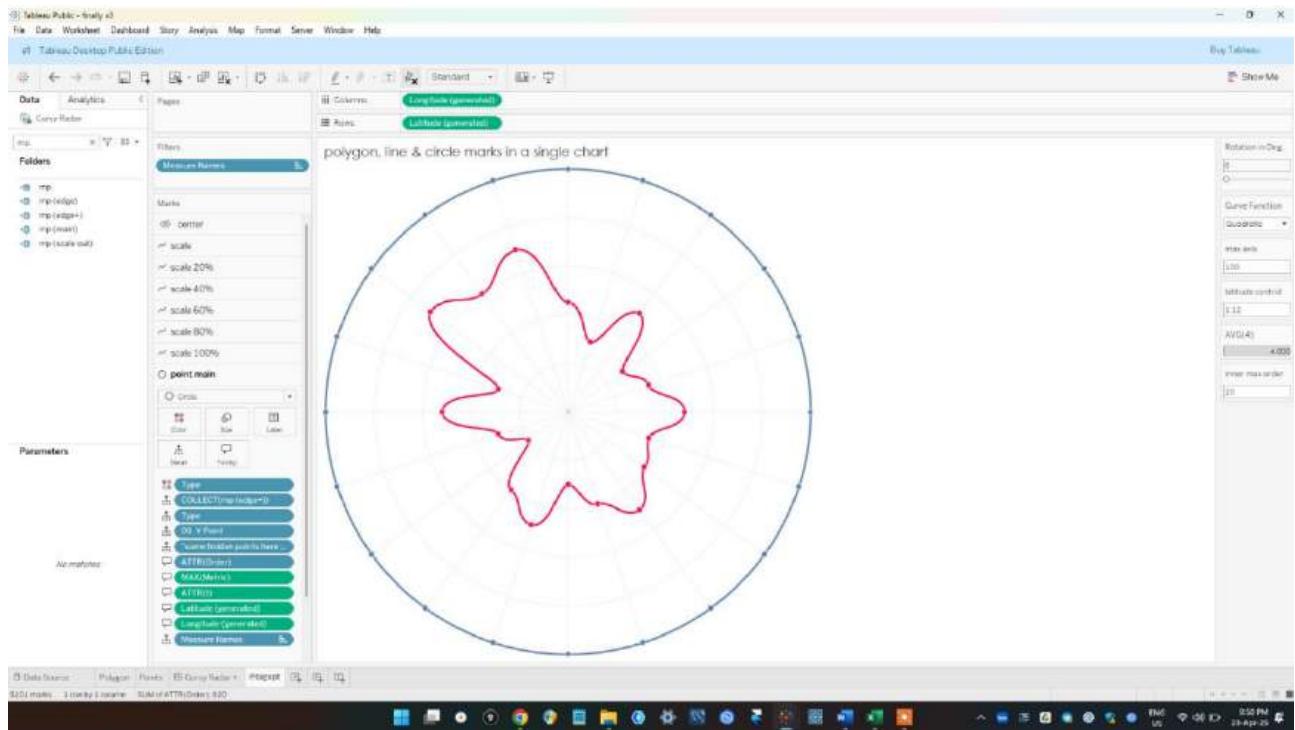
I was feeling excited after building this. Though it was not perfectly what I wanted but I knew I was getting close. I viewed the polygon style and got plenty questions and debugging to do.



First was, why was by polygon cutting the top and bottom of the circle? I spent over 2 hours trying to solve this, thinking it was a problem in my data or code settings. I completely forgot that this was a plot of latitude and longitude, and that latitude does not exceed 90°, so. How could you Seun!!! a whole 2 hours!!!



Well, there's still a lot to do, for the passion we don't give up...I fixed that and moved quickly onto building this.



Now that might feel like a massive leap in steps, but let me break it down in more detail.

Yes we no longer have a rectangular shape outside, it's now a circle.

This happened because I doubled the number of data points.

You see from Ken's template, there are 20 data points (I'll be using Radar 1 to explain this)

	A	B	C	D
1	Radar	Order	Metric	
2	Radar 1	1	45	
3	Radar 1	2	30	
4	Radar 1	3	50	
5	Radar 1	4	28	
6	Radar 1	5	35	
7	Radar 1	6	48	
8	Radar 1	7	35	
9	Radar 1	8	39	
10	Radar 1	9	50	
11	Radar 1	10	40	
12	Radar 1	11	30	
13	Radar 1	12	49	
14	Radar 1	13	40	
15	Radar 1	14	20	
16	Radar 1	15	30	
17	Radar 1	16	52	
18	Radar 1	17	30	
19	Radar 1	18	70	
20	Radar 1	19	60	
21	Radar 1	20	70	
22				
23				
24				
nc				

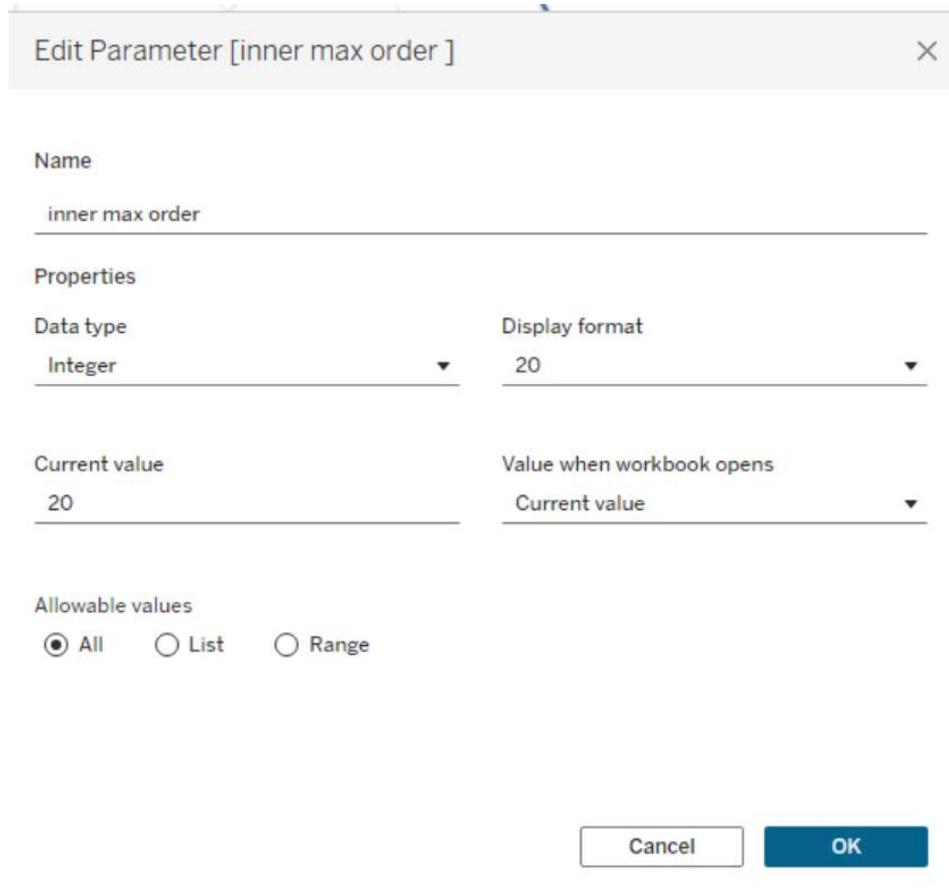
So, I doubled the data points, 20 for inner (for the red line), and 20 for outer (for the blue). He already created the ones for the inner, so the ones for the outer would be easy because there would be the same numbers which is the maximum data point in the circle (that's 100 in this case).

A	B	C	D	E
1	Type	Order	Metric	
12	Main	11	30	
13	Main	12	49	
14	Main	13	40	
15	Main	14	20	
16	Main	15	30	
17	Main	16	52	
18	Main	17	30	
19	Main	18	70	
20	Main	19	60	
21	Main	20	70	
22	edge	21	100	
23	edge	22	100	
24	edge	23	100	
25	edge	24	100	
26	edge	25	100	
27	edge	26	100	
28	edge	27	100	
29	edge	28	100	
30	edge	29	100	
31	edge	30	100	
32	edge	31	100	
33	edge	32	100	
34	edge	33	100	
35	edge	34	100	
36	edge	35	100	
37	edge	36	100	
38	edge	37	100	
39	edge	38	100	
40	edge	39	100	
41	edge	40	100	
42				

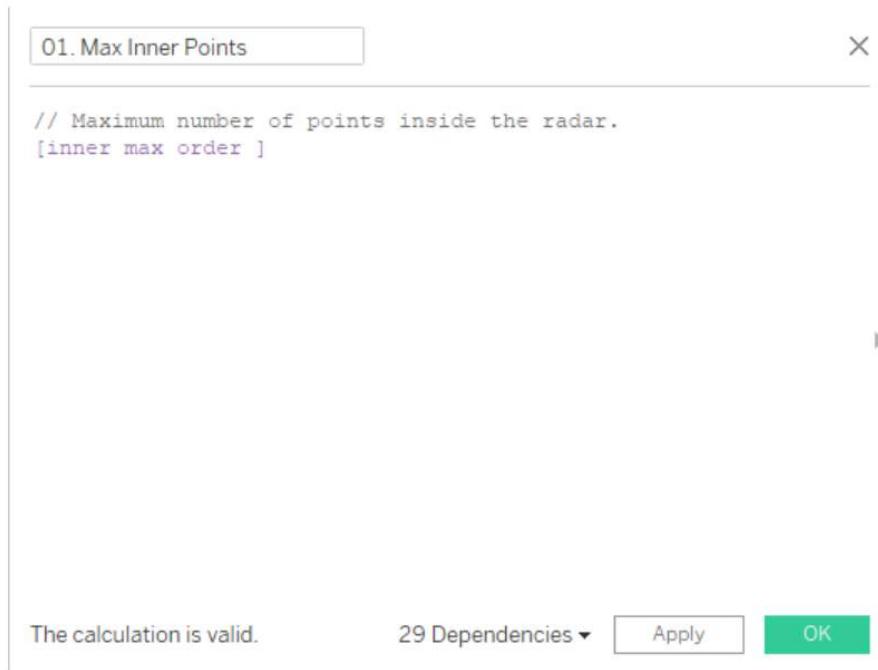
As you can see, there is an inner (Main) and the outer (edge). And for the Order 21 to 40 the values (metrics) are 100 which is the max value in this example.

This was what I used instead of the original template, and then, using the same model densification style, I adjusted some of the calculated fields.

First, I created an integer parameter and named it inner max order. This was the maximum order value in the inner polygon which is 20.



Next, I adjusted 01. Max Points I renamed it to 01. Max Inner Points and used this calculation



Next, I adjusted the calculated field 02. Next Row Metric

02. Next Row Metric

```
// Next row's metric comes from a join of the data to itself, getting the next row (Order + 1).
// The last row, therefore, won't have a value.
// This will get the value from the first row (Order = 1).
IF [Order] = [01. Max Inner Points]
THEN
    { FIXED : MAX(IF [Order] = 1 THEN [Metric] END) }

ELSEIF [Order] = {MAX([Order])}
THEN
    { FIXED : MAX(IF [Order] = [inner max order ] + 1 THEN [Metric] END) }

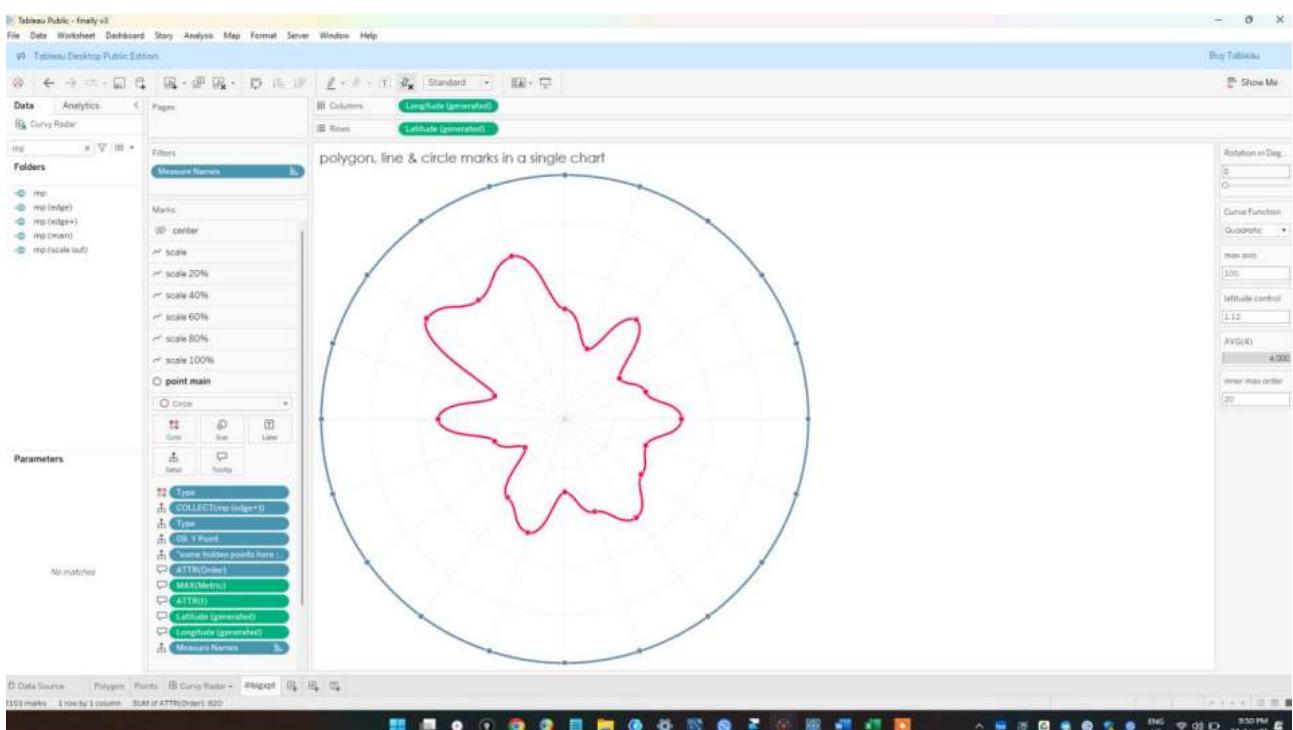
ELSE
    [Metric + Next]
END
```

The calculation is valid.

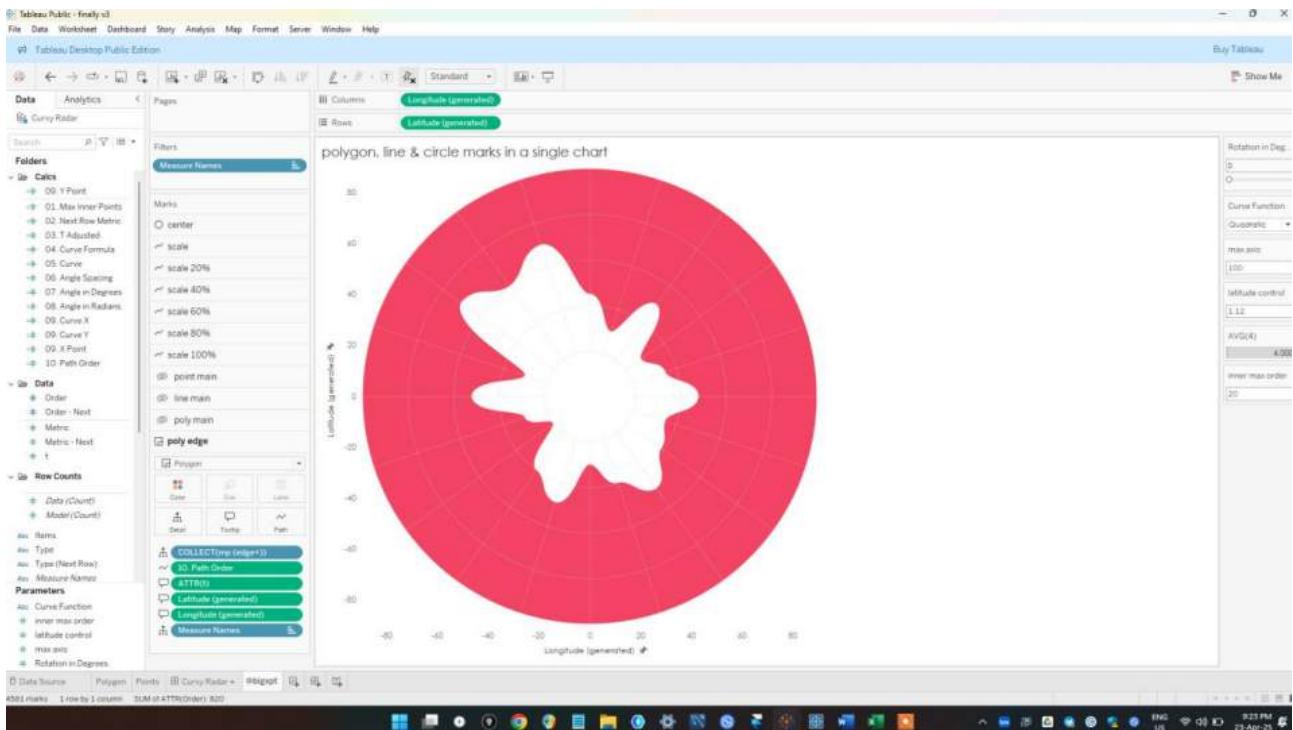
25 Dependencies ▾ Apply OK

This joins the last inner data point (order 20) to the first inner data point (order 1). Also, for the outer, max order (order 40), will be with joined to the first outer (order 21) that explains “inner max order + 1”.

So that's it! All the other calculated fields where left the same as Ken created them. So again, much credits to Ken. I just made a few changes and so we have this.



And after adjusting latitude and longitude so I don't waste another 2 hours, we have as a polygon this.



If you'd like to give it a go.... you can find the [template I made here](#). You will find I added a little bonus feature. To understand this better, just explore the idea and see what more you can develop from it! Enjoy.

CJ Round-up:

Wow I am seriously impressed with the creativity behind inverting the radar. I think this has major potential if used in a neat way like Seun's original visual to help layer different elements. Seun has built an incredible portfolio to date, and it has been amazing to watch his presence grow in the data community. **Make sure to check out his profile in full.**

LOGGING OFF,

CJ

## THE IRONVIZ CHAMPIONS SPORTS PORTFOLIO – BO MCCREADY

Hi all,

It is a great pleasure to have the 2025 Iron Viz Champion join the site this week. I'm sure you will be hearing and seeing a lot more of him in weeks to come, to talk all things competition, but I've managed to steal a moment of his time to talk about one of his passions, which is part of his growing sports portfolio.

CJ: Bo, great to have you join to share your interest in sports analytics. Before we dive into your portfolio – Let's start with a little about you – how did you get into data? How would you describe your public portfolio?

Bo: Thanks for having me here, CJ! My path into the data world was a little crooked, as I know many of ours are. I studied History as an undergraduate and didn't take a single math course. When I was working on my Master's in Public Policy, I was put into a required statistics course and I bought a copy of Moneyball to see if it would spark anything in me. Sure enough, it did, and I amazingly managed to do well in stats. I stayed in school through my Ph.D. and kept learning more about quantitative work before getting a job as a researcher with the local school district and starting to do data work full time. My public portfolio reflects my hobbies – you'll see lots of projects about sports, film, music, and history!

CJ: With a visual like your “**first round upset**”, you must have had so much match result data available it must have been hard to decide what metric to focus on? You landed on the first round upset results. How does finding that narrative within the data translate to then how you approached the chart choices and seed ordering?

# UPSET ALERT!

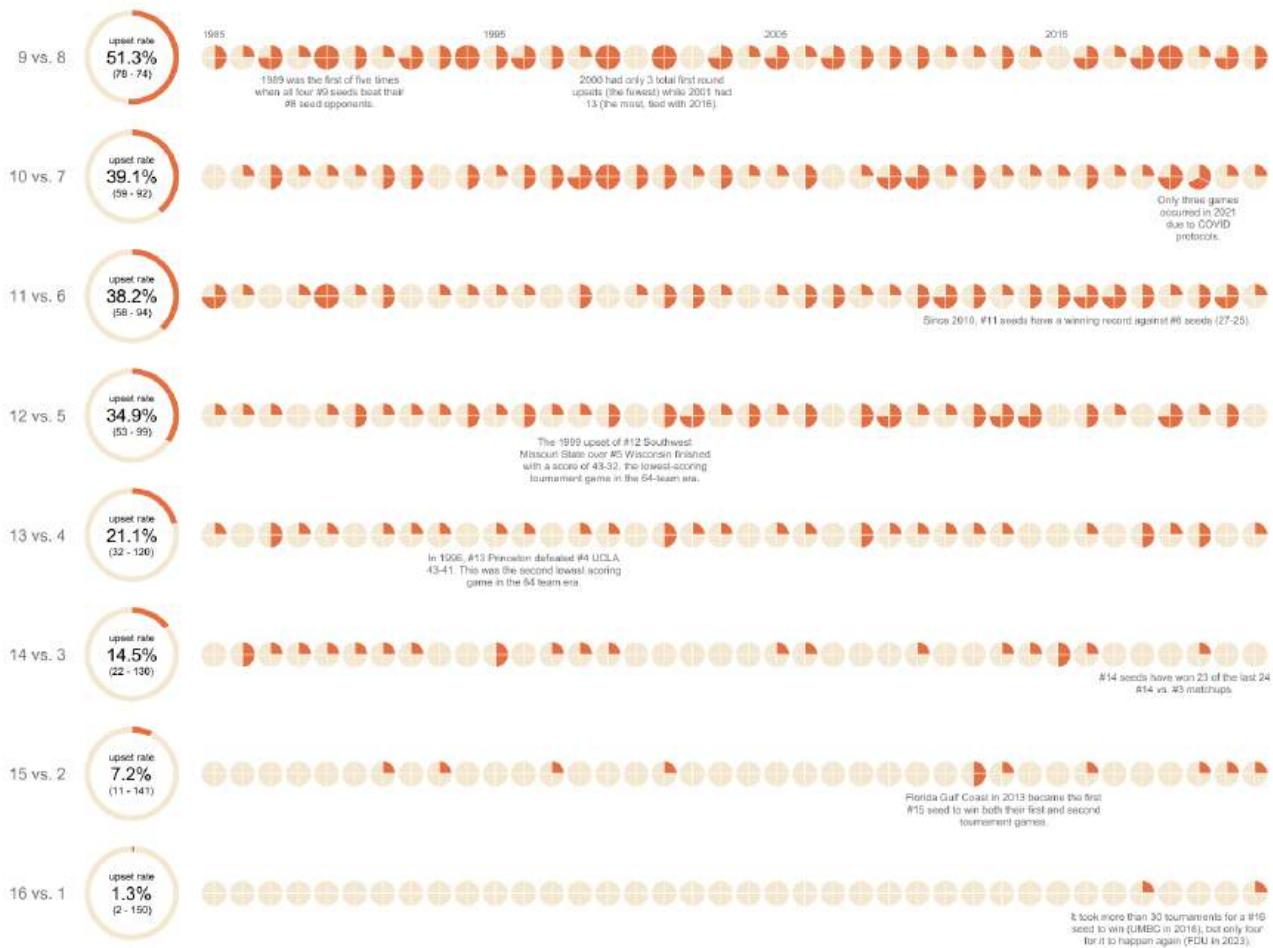
The history of first round upsets in the NCAA Men's Basketball Tournament

In 1985, the NCAA Men's Basketball Tournament switched to a format with 64 teams divided into four regions, with teams seeded from #1 (highest) to #16 (lowest) within each region.

These graphics show the results of each first round game since 1985, with upsets (lower seed beats higher seed) shaded in orange.

A project by Bo McCready  
@boknowsdata

Featuring data from  
Stathead/Sports Reference



Bo: In the USA, it's a ritual for so many people to fill out their NCAA tournament predictions and compete with their friends. We love debating which underdogs have a shot to win and those first round upsets tend to form so many people's opinions about whether a tournament is fun or not. So, as I was designing this project, I was imagining what someone might want to see when they were filling out their own predictions!

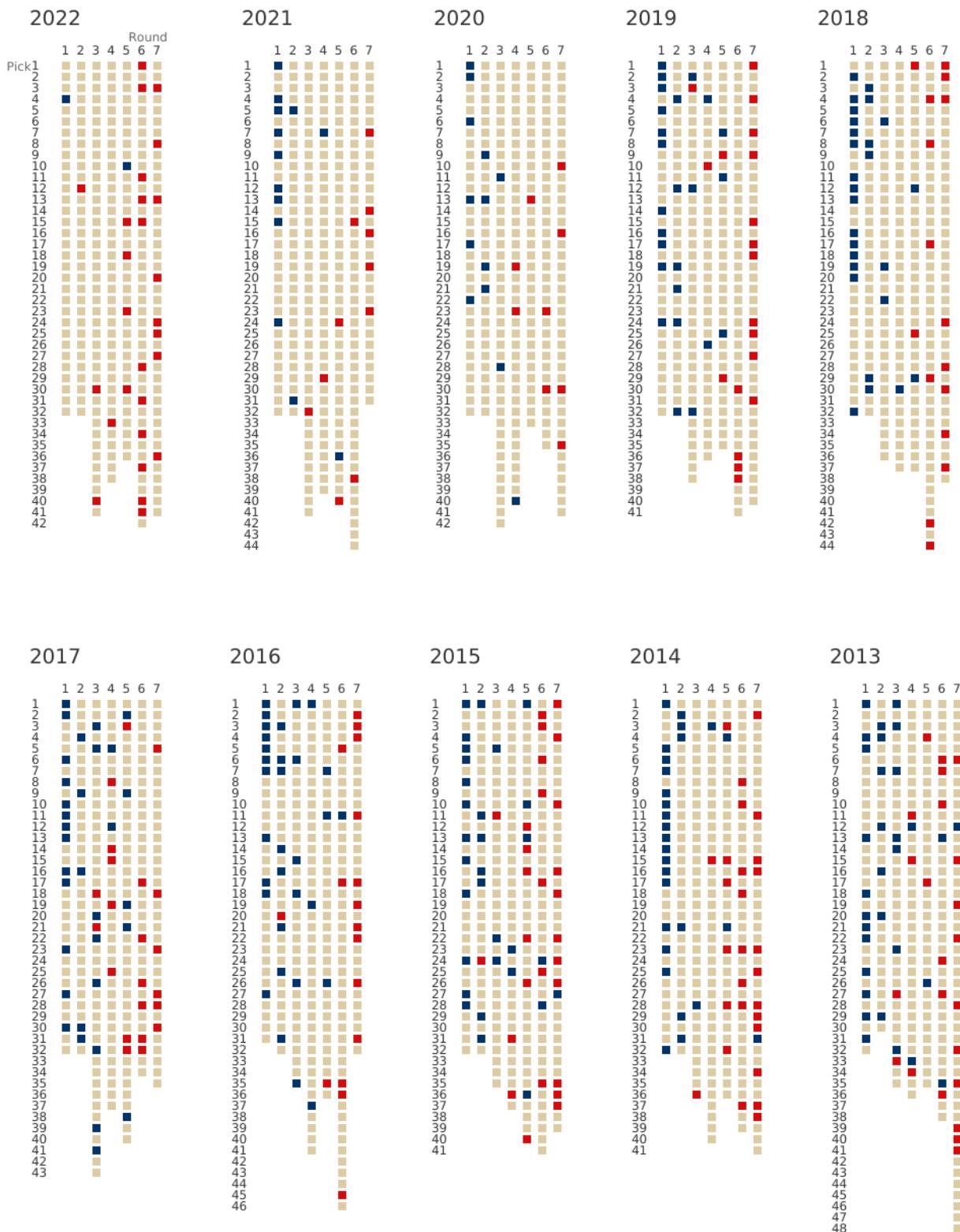
CJ: I love the column design to the **NFL draft picks** visually. There is something so appealing about the uniformity and reduced color highlighting. What do you think the drivers are behind the downturn in Pro Bowler first round picks and the upturn in picks with no NFL games in recent years?

# NFL Draft Picks 2013-2022

Designed by Bo McCready @boknowsdata  
Using data from Pro Football Reference

**Pro Bowler - NFL Player - No NFL Games**

Highlight Team  
All



Bo: I think there are a few factors coming into play. Some players might take a few years to actually make it into a game or to reach their first Pro

Bowl, so I bet we'd see a lot better representation as we move further into the future. I think we also see better strength and conditioning programs leading to longer careers, especially for the top athletes, so there may not be as many "spots" to go around.

CJ: The 2020 world cup “print” visual is a great way of breaking down the player information through group, team and club. The consistency in flags for domestic clubs for countries such as England, Qatar and Saudi really stand out from the page. Was this intentional?

In your opinion, does having players from multiple leagues/countries influence play style and chemistry?

Bo: Yes, the consistently absolutely is something I wanted to highlight! I think it's fascinating how basically all of the best English players in the world end up back home in the PL. It's the best league in the world and I think this project helped underline that for me. On the other hand, I do wonder how many nations who have their stars playing across the world can build chemistry in such a short time when preparing for something like the World Cup. I'd imagine managers of those more “global” teams would prioritize simpler tactics and allowing players to just be themselves on the pitch. I always felt that one of the reasons the USA always seemed to underperform their talent level with Gregg Berhalter was his rigid and complex system. I don't know if players who didn't have a lot of shared league experience could really handle that well.

# 2022 WORLD CUP SQUADS:

## PLAYER CLUB NATIONS

A PROJECT BY BO MCCREADY @BOKNOWSDATA

### GROUP A



### GROUP B



### GROUP C



### GROUP D



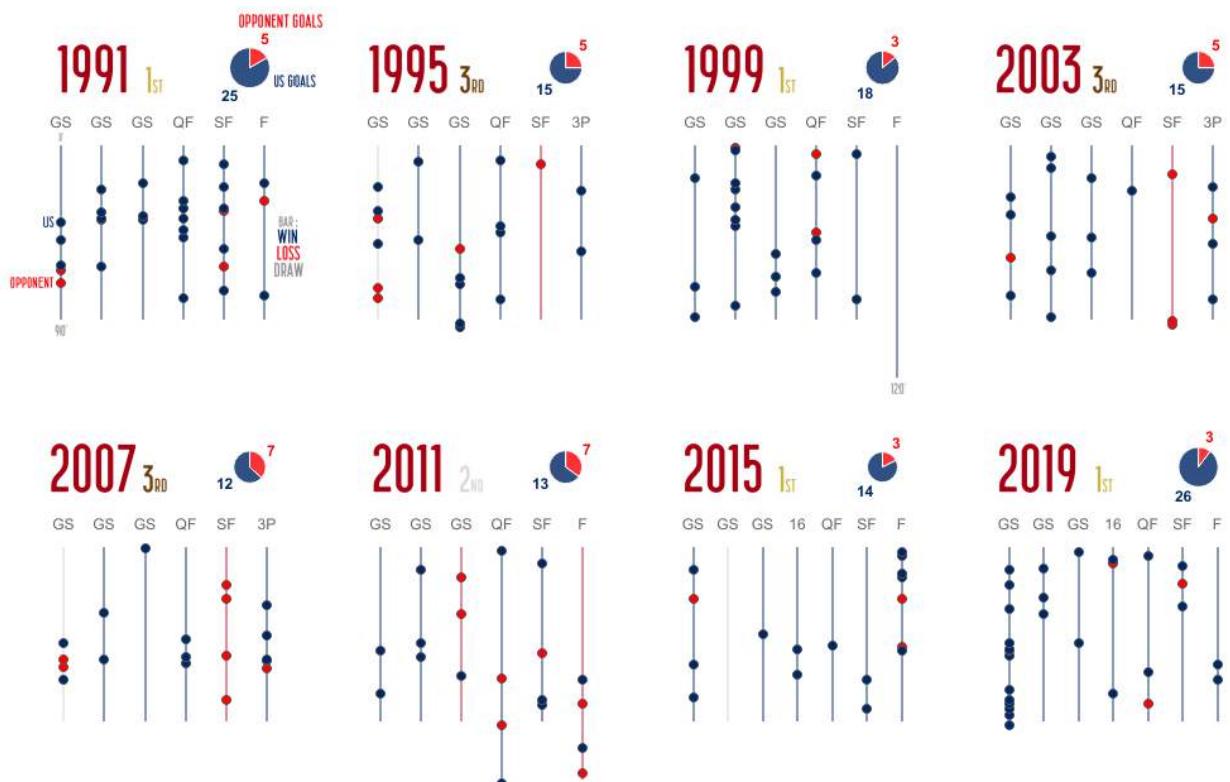
### GROUP E



CJ: Many of your visuals have a small multiple effect to them, and mirror some of your hobbies in prints. How do you make sure you can tell a story with just charts and no written context, such as in the USWNT World Cup visual?



## THE UNITED STATES WOMEN'S NATIONAL TEAM IN THE WORLD CUP



Visualization by Bo McCready @boknowsdata  
Data from fbref.com

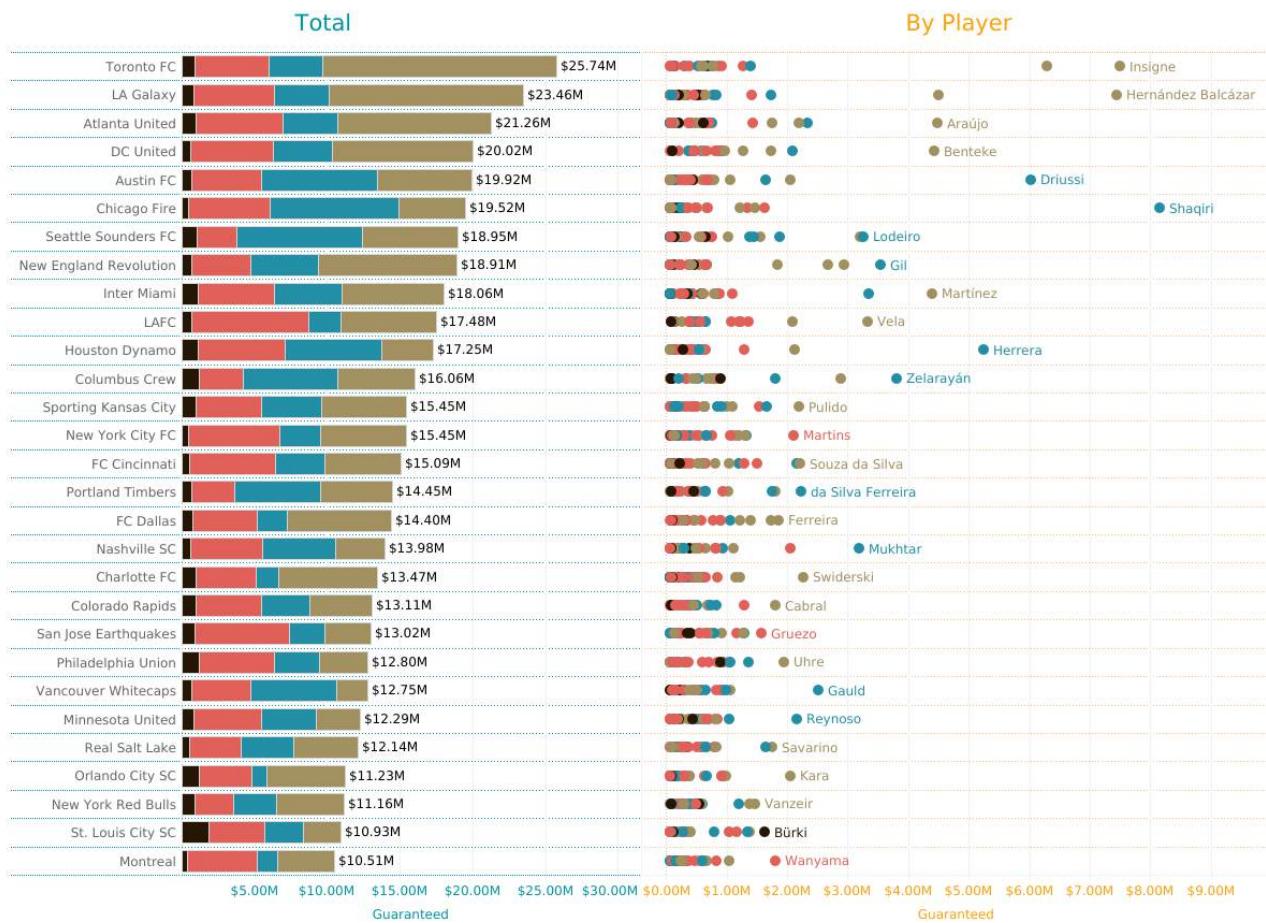
Bo: Telling a simple story with just charts is something that I love to try! DATA SOURCED FROM WIKIPEDIA AND ACCURATE AS OF 11/14/22 Since the start of my data viz journey, I have wanted to reach audiences who don't think about data as much as those of us who work with it all day. So I think it's always a fun challenge to think about accessible graphics with an easy, obvious message, and if they're understandable without written context, that's even better for audiences who don't speak English. It's always a thrill when I see a project of mine reshared with commentary in a language I don't understand, because that tells me the data and visuals were powerful enough to speak across the language divide.

CJ: Within your Major League Soccer Salary visual. What stood out to you in terms of spend composition amongst the different teams and

positions? How does the number of players in each position impact the ratio of total spend amongst positions?

## Major League Soccer Salaries May 2..

Position Filter  
All



Dashboard by Bo McCready @boknowsdata. Data comes from the official 5/16/23 MLSPA release.

Bo: Roster building in Major League Soccer is so strange and complex relative to the big European leagues! For me, it's always fascinating to see how teams choose to use their three Designated Player spots and which teams spend heavily on those spots. I'm a season ticket holder for Austin FC and this season, our team invested in attackers so that our entire attacking trio is Designated Players, with the rest of the squad relatively inexpensive. In general, I think you see MLS teams prioritizing attacking spending even with relatively fewer attackers in the squad because nothing really breaks a match open like a threatening attacker.

One thing you'll notice in this viz and in other MLS data is that the league has relative parity in investment, at least compared to European leagues. There are limitations in what you can pay your players, much like other major American sports leagues, and that system is designed to

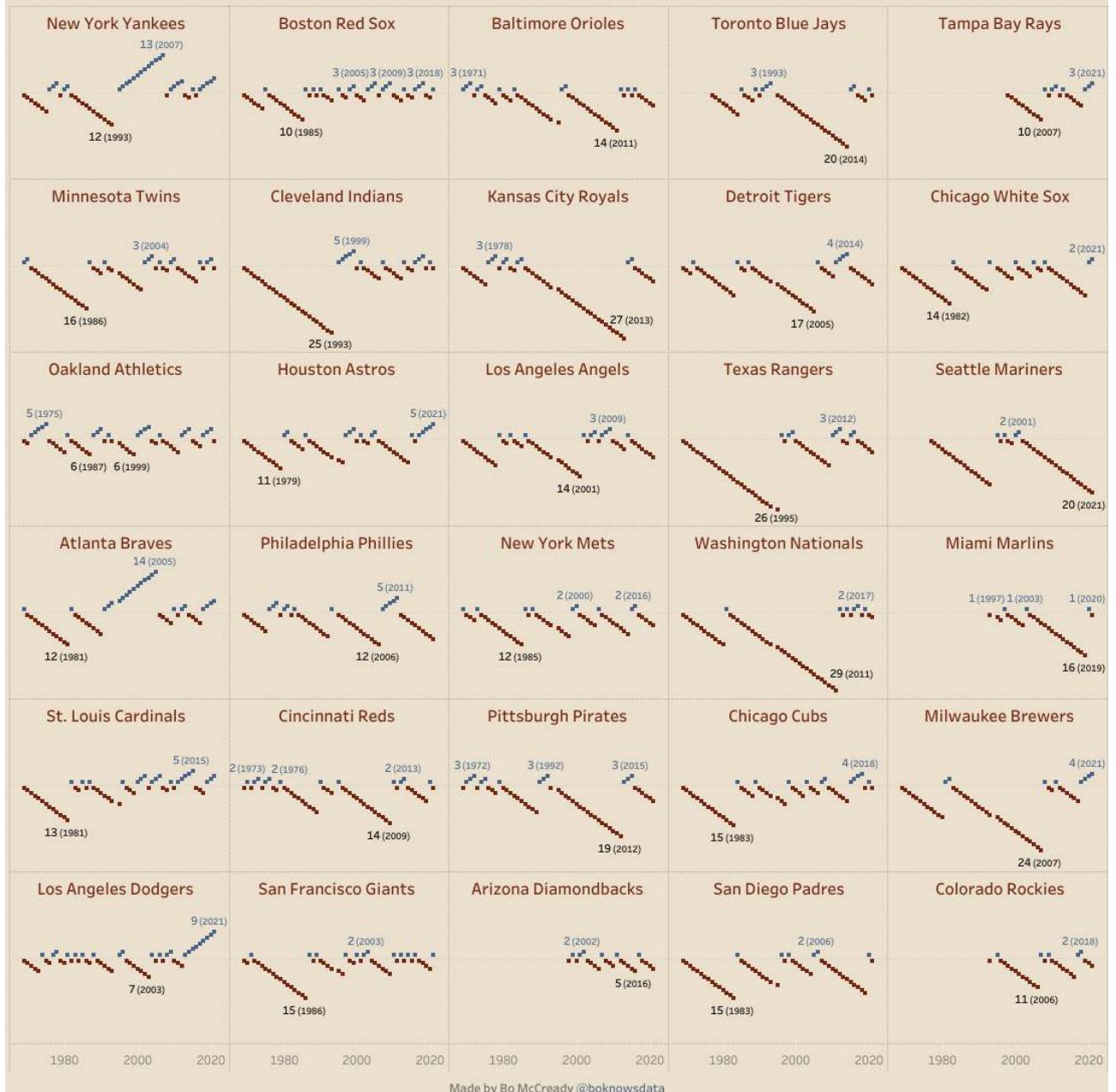
encourage consistent competition and avoid extended dominance by wealthy teams that just consolidate their position year after year. In that viz you linked, you'll notice Toronto outspending the bottom of the league at a 2.5:1 ratio, but that's nothing like comparing Man City and Ipswich, for example!

CJ: Your **MLB play off visual** emphasises the consistency of teams year on year making the playoffs. What impacted your choice to set the order of your small multiple in this way? What do you think the main drivers are behind teams like the Yankees of consistent play offs, vs Red Sox that have a more up down wave of streaks.

# PLAYOFF STREAKS AND DROUGHTS

*Major League Baseball, 1969-2021*

Graphs begin in 1969, when MLB added league championship series. Each franchise's longest streaks are labeled with their length and the year they ended.



Made by Bo McCready @boknowsdata

Bo: In this viz, each row is a division and then the teams are sorted from left to right by playoff appearances. I'm a Seattle Mariners fan and as you can see, that hasn't been very much fun. I think as much as people love to hate on the Yankees for their big spending, they also are just a very well run franchise. They spend big on their roster, but their player development is top notch and they are able to recruit free agents to play for the team as well as anyone. New York City is a huge draw and they capitalize on it well! The Red Sox, on the other hand, are still big spenders but don't tend to spend quite as much as the Yankees, so they

are more likely to have some downturns. And with the MLB playoff model, the Red Sox are competing against the other teams in the American League East, which are three lower budget but very smart franchises!

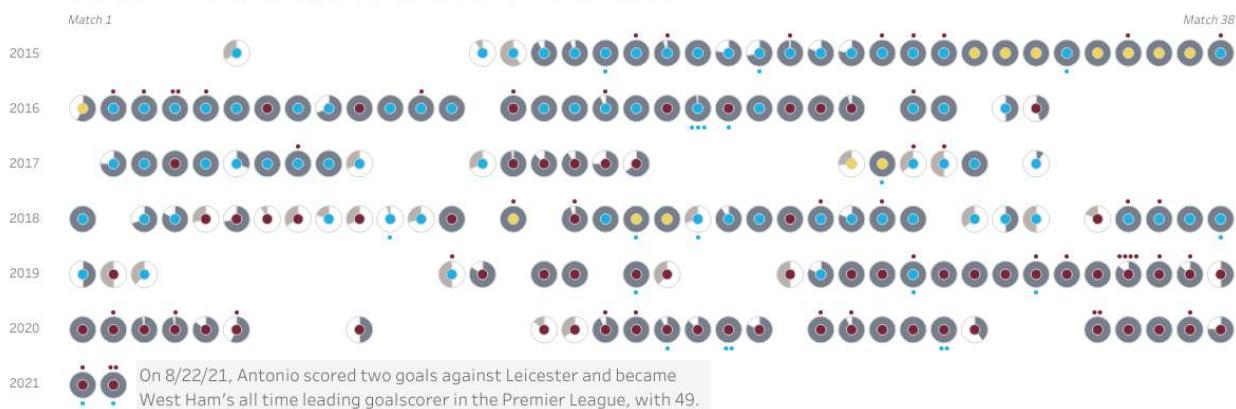
CJ: Great to see you throw in some West Ham visuals over the years. You created one that focussed on Michail Antonio's record breaking 49 all time premier league scorer for West Ham.

You are able to capture additional details around "blips" in season, and potential injury or fitness with little game time minutes in addition to the position shift. Antonio's goals – four against Norwich, 10 in total – were a key reason West Ham stayed up in 2019!

Is experimenting with combining different chart types (pie & circle shapes) as overlays something that has been of growing interest to you?

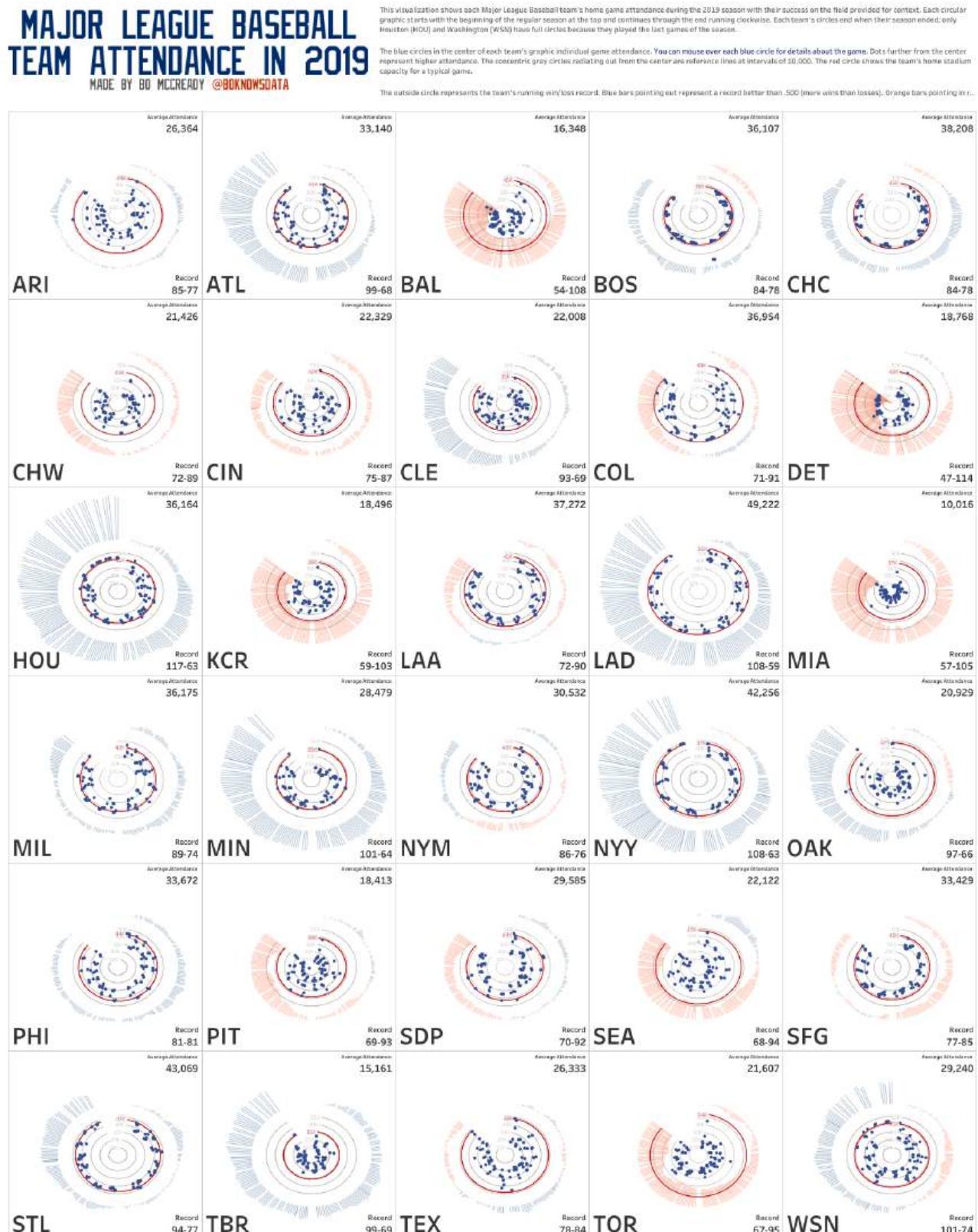
## MICHAIL ANTONIO WEST HAM'S RECORD SCORER

Each circle represents a Premier League match played by Michail Antonio for West Ham United. The filled portion of the outer circle represents minutes played as a **starter** or substitute. The inner circle shows whether Antonio started as a **defender**, **winger**, or **striker**. Dots above the circle represent goals scored and dots below represent assists.



Bo: Absolutely. I enjoy pushing boundaries and that includes using things like pies that often are frowned upon. I remember when Tableau didn't have floating transparent objects and ever since those became available, I've loved seeing what novel charts can be created simply by floating one thing over another. And I've used that style for multiple West Ham projects, including prints I made to commemorate their Europa Conference League championship. I was lucky enough to have one of those prints signed by members of the winning squad when they were in the USA last summer. Lukasz Fabianski called the project "really cool," and that's one of my favorite endorsements!

CJ: To close, let's talk about one of my favourites on your profile. The MLB Team 2019 attendance small multiple viz. What are some of the do's and don'ts when it comes to creating a radial viz that balances an interpretable custom chart type with design flair? What were some of the design changes you made as the viz developed?



Bo: This project was so much fun and even though it's been a few years, I remember the development process well! Honestly, the design didn't evolve much between the first ideas I had and the final project. I think the biggest addition along the way was adding stadium capacity for context. As for do's and don'ts, I guess I'd encourage experimenting with novel custom charts like this for data that is pretty conceptually simple. Although these charts are fancy, the basic idea of "how many people went to this baseball game" is straightforward. If I'm asking my end users to take on the mental load of understanding a novel chart type, I don't also want to ask them to take on the mental load of interpreting a complex metric. I would always suggest considering the complexity balance between a chart and the idea it's representing.

CJ Round-up: If you haven't had the opportunity to catch up on Iron Viz 2025 yet, check it out on [youtube here](#).

Once again, Thanks Bo for making time & congratulations on your success. Hope you had a wonderful conference. Looking forward to hearing and seeing more from Bo in future weeks.

LOGGING OFF,  
CJ

## FILTER SELECTIONS IN TABLEAU WITH ABHINAV

Hi all,

I've always enjoyed the 'design' aspect to visualisation tools. Anything related to user experience. The tableau community have been at the forefront of creating new ways to bend how Tableau works to try mirror good UX & UI principles when applied to their dashboards.

Today I'm delighted to be joined by Abhinav to walk us through his accordian filter selector. Over to you.



X LinkedIn

## Abhinav KR

Senior Tableau Developer @ Birdeye | Bengaluru, Karnataka, India

Welcome to my public portfolio.

[Read more](#)

[Follow](#)

[Hire Me](#)

Vizzes 10

Favorites 56

Following 14

Followers 207

The screenshot displays a grid of four Tableau visualizations (Vizzes) from Abhinav KR's portfolio:

- Gone Public in 2021**: Companies That Have Gone Public in 2021 | VOTD. Published by Abhinav KR. Rating: 4.8 (43 reviews), views: 18,564.
- Accelerators, Pulse, & Extensions in Action**: #TableauBuddyTUG - Supercharging Tableau.... Published by Abhinav KR. Rating: 0 (0 reviews), views: 70.
- Auto Sales & Trends | VOTD**: Published by Abhinav KR. Rating: 4.7 (125 reviews), views: 4,796.
- #MakeoverMonday | Harnessing Large Language Models (LLM)**: Published by Abhinav KR. Rating: 4 (4 reviews), views: 101.

I'm truly thankful for the opportunity to share this step-by-step walk through on your blog! In this post, I'll be diving into how I implemented an Accordion Menu in Tableau, combining it with a Pagination approach to enhance dashboard navigation and user experience. This technique helps create clean, structured layouts while maintaining interactivity, making dashboards more intuitive and engaging.

# Table Tennis

2008 Summer

## OLYMPIC HISTORY (1896 - 2016)

Athletes and Results

Hover over the bottom table to get more granular view ⓘ

Sport	Games	Hosting City	Events	NOCs	# Athletes	Medal	Gender	Age
Table Tennis	2008 Summer Olympics	Beijing	4	55	250	3%	50%	50%

Sport	Games	# Athletes	Gold	Male	Female	Under 30
Handball	2008 Summer Olympics	343	8%	50%	50%	78%
Hockey	2008 Summer Olympics	387	9%	51%	49%	86%
Judo	2008 Summer Olympics	386	4%	60%	40%	85%
Modern Pentathlon	2008 Summer Olympics	72	3%	50%	50%	85%
Rhythmic Gymnastics	2008 Summer Olympics					
Rhythmic Gymnastics	2008 Summer Olympics	95	7%		100%	100%
Rowing	2008 Summer Olympics	563	9%	65%	35%	83%

SELECTIONS

- Under 30
- Male
- Female
- Gold

MENU

- Medal ▾
- Gender ▾
  - Male
  - Female
- Age ▾
- Height ▾
- Weight ▾

Designed by: Abhinav KR | @kshorecabhinav

#LearnWithAbhinav

Data Viz Hub | @dsedata

Let me introduce myself, I'm Abhinav KR, a Senior Tableau Engineer at Birdeye with 8+ years of experience in Tableau, BI, and Data. I'm passionate about building dynamic and insightful visualizations, and I actively share Tableau tips, tricks, and tutorials on LinkedIn and YouTube through my platform, Data Viz Hub.

### Project Brief:

I came across a #B2VB (Back to Viz Basics) challenge in Tableau in 2022, where the theme was to create the best version of a text table. While exploring different approaches, I **found an insightful article by #InterWorks**, which introduced the concept of using an accordion menu in Tableau to optimize text table design.

Inspired by this, I decided to take it a step further by combining the Accordion Menu with a Pagination approach, allowing users to interact dynamically with large datasets while keeping the visualization clean and structured. Also taking steps even further I created a top table which gives you a preview of only the hover based selected row from the master table along with a extreme left sheet showing a dynamic big header for each sports.

# Handball

2008 Summer

## OLYMPIC HISTORY (1896 - 2016)

Athletes and Results

Hover over the liaison table to get more granular view ↗

Sport	Games	Hosting City	Events	NOCs	# Athletes	Medal	Gender	Age
Handball	2008 Summer Olympics	Beijing	2	16	343	85	50%	30%

Sport	Games	# Athletes	Gold	Male	Female	Under 30	SELECTIONS
Handball	2008 Summer Olympics	343	85	50%	50%	70%	<b>X Under 30</b>
Hockey	2008 Summer Olympics	307	95	51%	49%	60%	<b>X Male</b>
Judo	2008 Summer Olympics	306	45	62%	38%	65%	<b>X Female</b>
Modern Pentathlon	2008 Summer Olympics	72	35	50%	50%	65%	<b>X Gold</b>
Rhythmic Gymnastics	2008 Summer Olympics	30	79		100%	100%	
Rhythmic Gymnastics	2008 Summer Olympics	30	79		100%	100%	
Rowing	2008 Summer Olympics	563	99	65%	35%	85%	

Designed by: [danielkohli.com](#)

Published by: [danielkohli.com](#)

Hosted by: [danielkohli.com](#)

This method requires significant data preparation, including pivoting your data to work effectively with the accordion structure.

Proper calculations and parameter actions are essential to ensure seamless interactions.

If you'd like to follow along you can find the **DATA SET HERE**.

First we want to cleanse the data. I ended up identifying which of the 16 fields i needed, removing 5 fields and also removing any null values for

Age, Weight, Height, Medal etc.

Name	Gender	Age	Height	Weight	Region	Games	Hosting City	Sport	Medal
Juhamatti Tapio Aaltonen	M	28	184	85	Finland	2014 Winter	Sochi	Ice Hockey	Bronze
Paavo Johannes Aaltonen	M	28	175	64	Finland	1948 Summer	London	Gymnastics	Gold
Paavo Johannes Aaltonen	M	28	175	64	Finland	1948 Summer	London	Gymnastics	Gold
Paavo Johannes Aaltonen	M	28	175	64	Finland	1948 Summer	London	Gymnastics	Gold
Paavo Johannes Aaltonen	M	28	175	64	Finland	1948 Summer	London	Gymnastics	Bronze
Paavo Johannes Aaltonen	M	32	175	64	Finland	1952 Summer	Helsinki	Gymnastics	Bronze
Kjetil André Aamodt	M	20	178	85	Norway	1992 Winter	Albertville	Alpine Skiing	Bronze
Kjetil André Aamodt	M	22	178	85	Norway	1994 Winter	Lillehammer	Alpine Skiing	Bronze
Kjetil André Aamodt	M	22	178	85	Norway	1994 Winter	Lillehammer	Alpine Skiing	Silver
Kjetil André Aamodt	M	22	178	85	Norway	1996 Winter	Lillehammer	Alpine Skiing	Silver
Kjetil André Aamodt	M	28	178	85	Norway	1992 Winter	Albertville	Alpine Skiing	Gold
Kjetil André Aamodt	M	30	178	85	Norway	2002 Winter	Salt Lake City	Alpine Skiing	Gold
Kjetil André Aamodt	M	30	178	85	Norway	2002 Winter	Salt Lake City	Alpine Skiing	Gold
Kjetil André Aamodt	M	34	178	85	Norway	2006 Winter	Torino	Alpine Skiing	Gold
Pepijn Aan de Berg	M	26	189	72	Netherlands	1996 Summer	Atlanta	Rowing	Silver
Ann Kristin Aarne	F	23	182	64	Norway	1996 Summer	Atlanta	Football	Bronze
Giovanni Abagnale	M	21	198	98	Italy	2016 Summer	Rio de Janeiro	Rowing	Bronze
Jos Luis Abajo Gómez	M	35	194	87	Spain	2008 Summer	Beijing	Fencing	Bronze
Patimat Abakarova	F	21	165	49	Azerbaijan	2016 Summer	Rio de Janeiro	Taekwondo	Bronze
Manya Vasiliyevna Abakumova (-Tarskina)	F	22	179	88	Russia	2008 Summer	Beijing	Athletics	Silver
Aleksandr Aleksandrovich Abilmassov	M	28	180	83	Belarus	2008 Summer	Beijing	Canoing	Gold
Luc Abalo	M	31	182	86	France	2016 Summer	Rio de Janeiro	Handball	Silver

Some further cleaning: Group the fields like Age, Height & Weight into relevant buckets, and then removing the original fields.

Name	Gender	Age	Height	Weight	Region	Games	Hosting City	Sport	Medal
Juhamatti Tapio Aaltonen	M	Under 30	181-200 cm	71 - 140 kg	Finland	2014 Winter	Sочи	Ice Hockey	Bronze
Päivö Johannes Aaltonen	M	Under 30	181-190 cm	Under 70 kg	Finland	1948 Summer	London	Gymnastics	Gold
Päivö Johannes Aaltonen	M	Under 30	181-190 cm	Under 70 kg	Finland	1948 Summer	London	Gymnastics	Gold
Päivö Johannes Aaltonen	M	Under 30	181-190 cm	Under 70 kg	Finland	1948 Summer	London	Gymnastics	Gold
Päivö Johannes Aaltonen	M	Under 30	181-190 cm	Under 70 kg	Finland	1948 Summer	London	Gymnastics	Bronze
Päivö Johannes Aaltonen	M	Under 30	181-190 cm	Under 70 kg	Finland	1948 Summer	London	Gymnastics	Bronze
Kjetil André Aamodt	M	Under 30	181-190 cm	71 - 140 kg	Norway	1992 Winter	Albertville	Alpine Skiing	Bronze
Kjetil André Aamodt	M	Under 30	181-190 cm	71 - 140 kg	Norway	1994 Winter	Lillehammer	Alpine Skiing	Bronze
Kjetil André Aamodt	M	Under 30	181-190 cm	71 - 140 kg	Norway	1994 Winter	Lillehammer	Alpine Skiing	Silver
Kjetil André Aamodt	M	Under 30	181-190 cm	71 - 140 kg	Norway	1994 Winter	Lillehammer	Alpine Skiing	Silver
Kjetil André Aamodt	M	Under 30	181-190 cm	71 - 140 kg	Norway	1992 Winter	Albertville	Alpine Skiing	Gold
Kjetil André Aamodt	M	Under 30	181-190 cm	71 - 140 kg	Norway	2002 Winter	Salt Lake City	Alpine Skiing	Gold

When we connect to it in Tableau we will need to pivot the data. This is in order for us to enable our end users to select Fields in our dashboard.

The screenshot shows the Tableau Data Editor interface. A context menu is open over the last row of the data table. The menu items are: 'Rename', 'Copy Values', 'Hide', 'Create Calculated Field...', 'Pivot' (highlighted in blue), and 'Merge Mismatched Fields'.

Name	Gender	Age	Height	Weight	Region	Games	Hosting City	Sport	Medal
Juhamatti Tapio Aaltonen	M	Under 30	181-200 cm	71 - 140 kg	Finland	2014 Winter	Сочи	Ice Hockey	Bronze
Päivö Johannes Aaltonen	M	Under 30	181-190 cm	Under 70 kg	Finland	1948 Summer	London	Gymnastics	Gold
Päivö Johannes Aaltonen	M	Under 30	181-190 cm	Under 70 kg	Finland	1948 Summer	London	Gymnastics	Gold
Päivö Johannes Aaltonen	M	Under 30	181-190 cm	Under 70 kg	Finland	1948 Summer	London	Gymnastics	Gold
Päivö Johannes Aaltonen	M	Under 30	181-190 cm	Under 70 kg	Finland	1948 Summer	London	Gymnastics	Bronze
Päivö Johannes Aaltonen	M	Under 30	181-190 cm	Under 70 kg	Finland	1948 Summer	London	Gymnastics	Bronze
Kjetil André Aamodt	M	Under 30	181-190 cm	71 - 140 kg	Norway	1992 Winter	Albertville	Alpine Skiing	Bronze
Kjetil André Aamodt	M	Under 30	181-190 cm	71 - 140 kg	Norway	1994 Winter	Lillehammer	Alpine Skiing	Bronze
Kjetil André Aamodt	M	Under 30	181-190 cm	71 - 140 kg	Norway	1994 Winter	Lillehammer	Alpine Skiing	Silver
Kjetil André Aamodt	M	Under 30	181-190 cm	71 - 140 kg	Norway	1992 Winter	Albertville	Alpine Skiing	Gold
Kjetil André Aamodt	M	Under 30	181-190 cm	71 - 140 kg	Norway	2002 Winter	Salt Lake City	Alpine Skiing	Gold

Also create a metric value for the unique number of Athletes/Players as a new calculated field //Athletes/Players

COUNTD(Name)

Now we are ready to build.

Building the parameter for the Sidebar Header having a toggle option

Edit Parameter [Sidebar Filter (P)] X

**Name**  
Sidebar Filter (P)

**Properties**

<b>Data type</b> <input type="text" value="Integer"/>	<b>Display format</b> <input type="text" value="2"/>
<b>Current value</b> <input type="text" value="2"/>	<b>Value when workbook opens</b> <input type="text" value="Current value"/>

**Allowable values**

All    
  List    
  Range

Value	Display As
0	0
1	1
2	2
3	3
4	4
5	5
Click to add	

Fixed  
 When workbook opens  
[Add values from ▾](#)

Remove Selected
OK
Cancel

Create calculated fields for Age, Gender, Height, Medal & Weight so that we can have the option of Adding or removing these fields in our dashboard.

Define a calculated field that switches between 0 and 1:

Sidebar Calc for Age:

IF = 1

THEN 0

ELSE 1

END

Follow these steps for the remaining four fields like Gender, Height, Weight & Medal.

Sidebar Value - Weight

#B2VB\_W03\_Table

```
IF [Sidebar Filter (P)] = 5  
THEN 0  
ELSE 5  
END
```

Creating the drop down / Collapse toggle leveraging shapes  
Assign a down arrow shape to 0 (collapsed) and an up arrow to 1  
(expanded) for Age.

See the example below when the parameter is having a value of 0:

The screenshot shows the Tableau Data Source pane. In the 'Fields' section, there is a single field named 'Age Toggle'. Below it, under 'Marks', is a 'Shape' card with several options: Color, Size, Label, Detail, Text, and Shape. The 'Shape' dropdown is expanded, showing a small downward-pointing triangle icon. A tooltip for this icon reads 'MIN([Sidebar Value - Age])'. In the 'Sidebar Filter (P)' dropdown, the value '0' is selected. At the bottom of the pane, the expression 'MIN([Sidebar Value - Age])' is displayed.

Age is being shown as a label along with the downwards arrow means it is collapsed.

Similarly when the parameter has a value of 1, it shows as the expanded window.

This screenshot is identical to the one above, except the 'Sidebar Filter (P)' dropdown now contains the value '1'. This change causes the shape to expand, revealing a larger upward-pointing triangle icon. The tooltip for this icon also remains 'MIN([Sidebar Value - Age])'.

Repeat the same process for the Height, Weight, Gender & Medal and assign:

a down arrow shape to 0 (collapsed) and an up arrow to 2 (expanded) for Gender.

a down arrow shape to 0 (collapsed) and an up arrow to 3 (expanded) for Medal.

a down arrow shape to 0 (collapsed) and an up arrow to 4 (expanded) for Height.

a down arrow shape to 0 (collapsed) and an up arrow to 5 (expanded) for Weight.

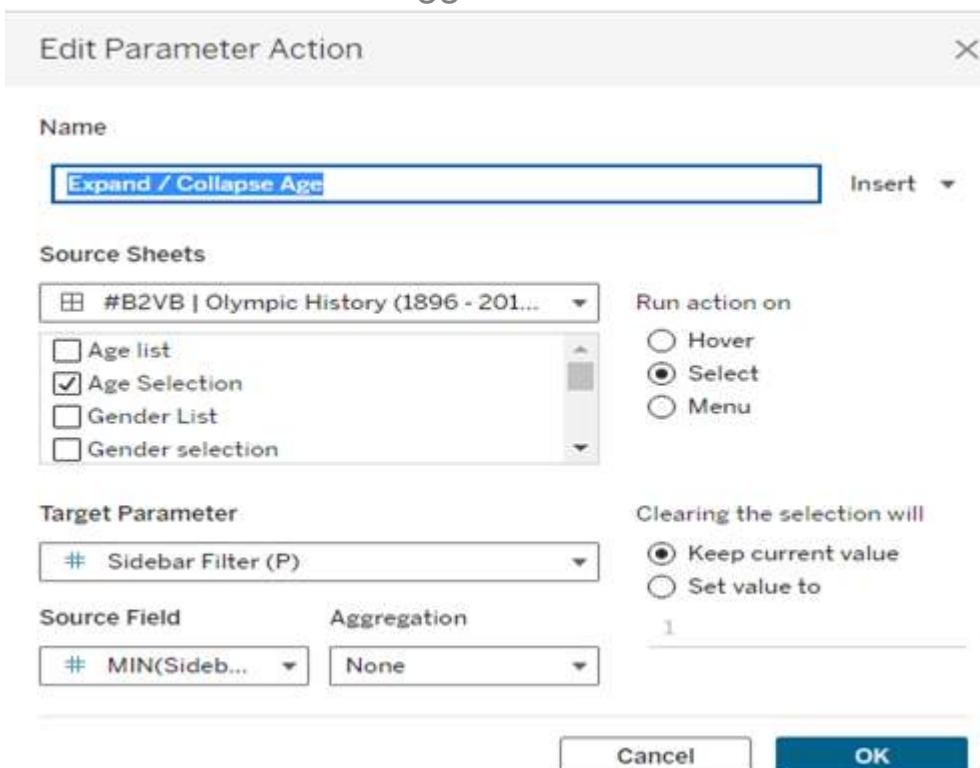
Adding dashboard level interactivity for the toggle items

Adding a parameter action that triggers the parameter when the toggle button is clicked.

Clicking the collapsed icon (1) returns 0, which opens the menu.

Clicking the expanded icon (0) sets it to 1, which collapses the menu

For example, let's see how we are creating the parameter action for Age Toggle button:



Ensure we do the same for Height, Weight & Medal.

You will need to create two calculated fields to avoid getting the icons getting selected once clicked for a smooth UX.

Create two calculated fields TRUE & FALSE and add them to the Details shelf in the Marks pane.

False #B2VB\_W03\_Table X

FALSE

The calculation is valid. 16 Dependencies ▾ Apply OK

Marks

Shape

Color	Size	Label
Detail	Tooltip	Shape

MIN(Sidebar Value - Gender)

True

False

Add individual filter action for each of the five toggle selection buttons.

Edit Filter Action

Name: Deselect Gender Header Insert ▾

**Source Sheets**

#B2VB | Olympic History (1896 - 201... ▼

- Age list
- Age Selection
- Gender List
- Gender selection

**Run action on**

- Hover
- Select
- Menu
- Single-select only

**Target Sheets**

Gender selection ▼

**Clearing the selection will**

- Keep filtered values
- Show all values
- Exclude all values

#### Filter

All fields     Selected fields

	Source Field	Target Data Source	Target Field
<input type="checkbox"/>	True	#B2VB_W03_Table	False
<input type="checkbox"/>	Click to add		

Creating the selection List to be displayed once the toggle expands.

Starting with the Age as an example here.

We are basically going to create a bar chart with a placeholder of MIN(1) and then will add the Field Name equals Age on filter and use Field Value on the Rows shelf. Also make sure to restrict the axis to start at 0 & end at 1.

The screenshot shows the Tableau Data Source interface. In the 'Fields' section, 'Age' is selected. In the 'Columns' section, 'AGG(MIN(1))' is selected. In the 'Rows' section, 'Field Value' is selected. The 'Age List' section displays categories: Under 30, 31-50, 51-70, and 70+.

Right click on the Field Value and create a set named as Selected Field Value.

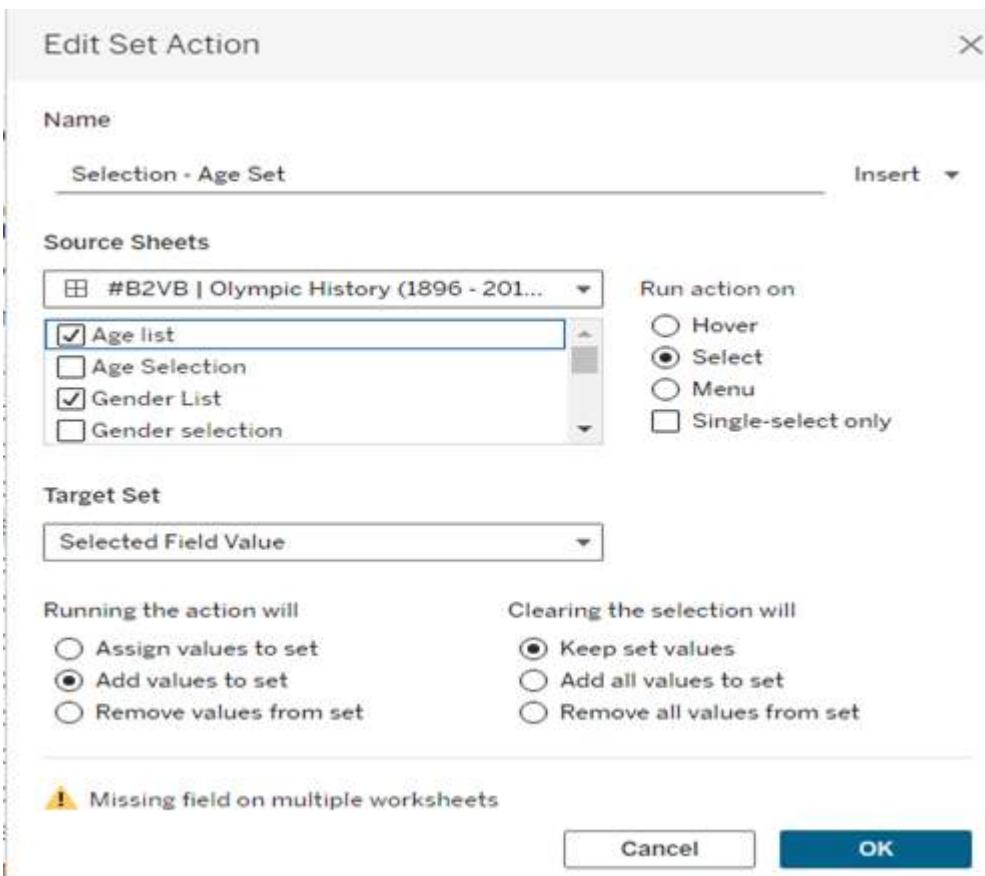
Now use this set on color and provide the relevant colors for the IN & OUT options in the set.

The screenshot shows the Tableau Data Source interface with the 'Selected Field Value' set applied to the 'Field Value' field. The 'Age list' section now shows the categories colored according to the set.

Repeat the same steps for the rest of the toggle items like Gender, Height, Weight & Medals to have their own selection lists.

Let's create a set action, the selection of the fields in the list.

This step ensures that once clicked on the list it adds the field value.



Create the set action as shown above and select all the sheets i.e. Age List, Gender List, Height List, Weight List & Medal List in the source sheets and select the Selected Field Value set in the Target Set also select Running the action will Add values to set.

Let's create the Selection List that should be present as the output for whatever fields are added or removed as shown below:

Ensure that we are using the same set Selected Field Value to create this chart.

Let's create the Removal Set for the Selection List:

Edit Set Action X

Name: Remove - Age Set Insert ▾

Source Sheets: #B2VB | Olympic History (1896 - 201... Run action on

Selection list ○ Hover  
 Sheet 17 ● Select  
 Table ○ Menu  
 Weight List □ Single-select only

Target Set: Selected Field Value

Running the action will:  Assign values to set Clearing the selection will  
 Add values to set  Keep set values  
 Remove values from set  Add all values to set  
 Remove all values from set

In the source sheet, use only the Selection List and in the Target Set use the Selected Field Value and under the Running the action will, use the Remove values from set.

At this step we are done with the custom selection/removal option in the dashboard.

## BUILDING THE TABLE

Creating a dynamic text table that updates based on the selection

Create a LOD calc for the Total Number of Athletes:

{FIXED , : SUM(IF = 'Age Range' THEN END)}  
//Fixed by Games & sports

Then create another calc for the % of Athletes as:

SUM() / SUM()

Create the text table as shown below:

Table (2) Data:

Sport	Games	# Athletes	Gold	Male	31-50	161-180 cm
Aeronautics	1936 Summer Olympics	1	100%	100%		
Alpine Skiing	2014 Winter Olympics	687	2%	56%	12%	68%
	2010 Winter Olympics	685	1%	56%	11%	66%
	2006 Winter Olympics	619	2%	57%	7%	73%
	2002 Winter Olympics	559	2%	56%	6%	74%
	1998 Winter Olympics	477	2%	53%	4%	73%
	1994 Winter Olympics	544	2%	55%	1%	70%
	1992 Winter Olympics	750	1%	65%	2%	47%
	1988 Winter Olympics	666	2%	64%	2%	43%
	1984 Winter Olympics	390	2%	69%	0%	66%
	1980 Winter Olympics	327	2%	62%	1%	57%
	1976 Winter Olympics	381	2%	70%	0%	44%
	1972 Winter Olympics	322	2%	62%	3%	82%
	1968 Winter Olympics	407	1%	71%		81%
	1964 Winter Olympics	399	2%	69%	4%	57%
	1960 Winter Olympics	317	2%	60%	1%	25%
	1956 Winter Olympics	383	2%	68%	6%	16%
	1952 Winter Olympics	378	2%	66%	6%	11%
	1948 Winter Olympics	357	2%	75%	5%	11%

We are only left with the upper table that shows the filtered preview based on hovering the master table:  
Create a new calc as Sort Options:

```
IF THEN
ELSE 'Athletes'
END
```

And use this to create a Sort Options Set using this and now create the Upper preview table as shown below:

Table Data:

Sport	Games	Hosting City	Events	NOCs	# Athletes	Median	Gender	Age	Height
Aeronautics	1936 Summer Olympics	Berlin	1	1	1	100%	100%		
Alpine Skiing	1994 Winter Olympics	Garmisch-Partenkirchen	2	26	222	2%	64%	39	169
	1998 Winter Olympics	Saint Moritz	6	39	367	3%	70%	39	176
	2002 Winter Olympics	Utah	8	38	378	3%	66%	39	176
	2006 Winter Olympics	Cortina d'Ampezzo	8	29	382	2%	69%	39	180

And now just add the dashboard filter action on this dashboard based on the master table.

Almost finished!

Arrange the dashboard items, sheets & container as shown in the below format

## OLYMPIC HISTORY (1896 - 2016)

Athletes and Results

Hover over the bottom table to get more granular view 

Sport	Games	Hosting City	Events	NOCs	# Athletes	Age
Triathlon	2008 Summer Olympics	Beijing	2	26	110	30%



The screenshot shows a data visualization interface. On the left is a main table with data for various sports at the 2008 Summer Olympics. On the right is a sidebar with filter controls.

Sport	Games	# Athletes	Gold	Male	31-50	161-180 cm
Swimming	2008 Summer Olympics	1,749	4%	51%	7%	49%
Synchronized Swimming	2008 Summer Olympics	117	9%		6%	91%
Table Tennis	2008 Summer Olympics	250	3%	50%	27%	63%
Taekwondo	2008 Summer Olympics	125	6%	50%	8%	62%
Tennis	2008 Summer Olympics	254	2%	50%	16%	53%
Trampolining	2008 Summer Olympics	32	6%	50%	13%	75%
Triathlon	2008 Summer Olympics	110	2%	50%	30%	62%

**SELECTIONS**

- 31-50
- Male
- 161-180 cm
- Gold

**MENU**

- Medal
  - Gold
  - Silver
  - Bronze
- Gender
- Age
- Height
- Weight

To enhance the UX and to avoid having a scroll option in the table, we should be using the pagination approach to let the end users switch between the different pages.

Create Parameters:

Page Number (Integer, starts at 1).

Rows per Page (Integer, e.g., 10, 20, 50).

Create Calculated Fields:

Index Calculation: INDEX()

Pagination Filter:

IF INDEX() > ( \* ( - 1))

AND INDEX() <= ( \* )

THEN "Show" ELSE "Hide" END

Previous Page:

IF > 1 THEN – 1 ELSE END

Next Page:

IF < CEILING(COUNTD() / )

THEN + 1 ELSE END

Page Numbers Preview: STR(INDEX())

Apply Filters & Add Controls:

Filter Pagination Filter → Select "Show".

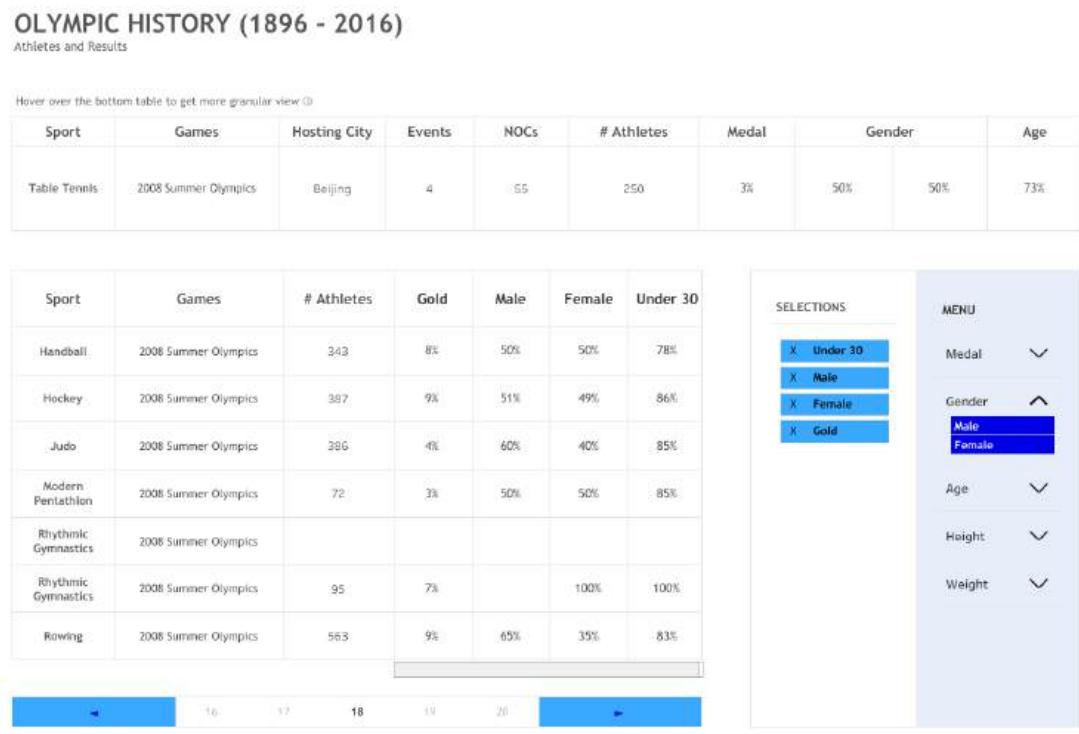
Left Arrow (Shape Button) → Set Page Number to .

Right Arrow (Shape Button) → Set Page Number to .

Page Preview (Clickable Page Numbers) → Use Page Numbers  
 Preview with Parameter Action.  
 Deselect Option:  
 Use the same TRUE FALSE fields and create a filter action for all the three sheets, Left & Right Arrow icon, and the preview page number window individually.  
 And that's it, just clean the dashboard, tweak the formatting and **we are done.**

# Table Tennis

2008 Summer



Designed by: Abhinav KR | @kishoreabhinav

#LearnWithAbhinav

Data Viz Hub | @dsedata

This project was an exciting challenge that pushed me to explore creative ways to improve text table design in Tableau. By leveraging accordion menus and pagination, I was able to create a solution that balances data density with interactivity, making it easier for users to navigate large datasets without overwhelming them.

I hope this walk-through helps others looking to implement better UX in Tableau dashboards. Looking forward to sharing this with the community and hearing everyone's thoughts!

Thanks,  
Abhinav KR

CJ Round-up: I love the idea of having the filters visible on the screen. It has a great business user case often where a report might get prinscreened its always good to know whats actually been included. The pagination and accordion really help make it a cohesive piece and complement the filter selection so nicely.

LOGGING OFF,  
CJ

## INTRO TO GITHUB PAGES & HTML

Hi all,

Have you ever wanted a unique way to share your code, projects, repositorys and portfolio?

How about wanting a “website” but not necessarily wanting to pay the maintenance and hosting fees?

Well today we will look at how you can set up with a github HTML template to showcase your work.

First you'll want to log in to Github.

The screenshot shows the GitHub Home page. At the top, there's a search bar with 'Find a repository...' and a 'New' button. Below the search bar, there's a section for 'Chelsea-Fc-Performance-Insights/Competition'. The main area features several cards: 'Learn with a tutorial project' (Introduction to GitHub, Code with Copilot), 'GitHub Pages' (Create a site or blog from your GitHub repositories with GitHub Pages), and 'Hello GitHub Actions' (Create a GitHub Action and use it in a workflow). Below these, there's a link to 'See more tutorial projects'. On the left, there's a sidebar with 'Top repositories' and a 'New repository' button. On the right, there's a 'Start writing code' section and a 'Create a new repository' form. The 'Create a new repository' form includes fields for 'Repository name' (with placeholder 'name your new repository...'), 'Visibility' (Public vs. Private), and a 'Create' button.

& we can straight away make a repository. You'll need to make it public.  
That's okay though as for now it doesn't have anything in it!

The screenshot shows the 'Create a new repository' form. It has a title 'Create a new repository' and a sub-instruction 'A repository contains all your project files, including the revision history. Already have a project repository elsewhere? Import a repository.' Below this, it says 'Required fields are marked with an asterisk (\*).'. There's a 'Repository template' dropdown set to 'No template'. A note says 'Start your repository with a template repository's contents.' Under 'Owner \*', it shows 'CJ-Mayes' and 'Repository name \*' with 'Portfolio' entered. A note says 'Great repository names are short and memorable. Need inspiration? How about iterate-adventure ?'. There's a 'Description (optional)' field with a placeholder ' '. Below this, there's a visibility choice between 'Public' and 'Private', with 'Public' selected. A note says 'Anyone on the internet can see this repository. You choose who can commit.' Below that, there's a note for 'Private' repositories: 'You choose who can see and commit to this repository.'

For this run through we will create a HTML site. Google “Free HTML Templates” .... there are a ridiculous amounts online.

Here are a few sites:

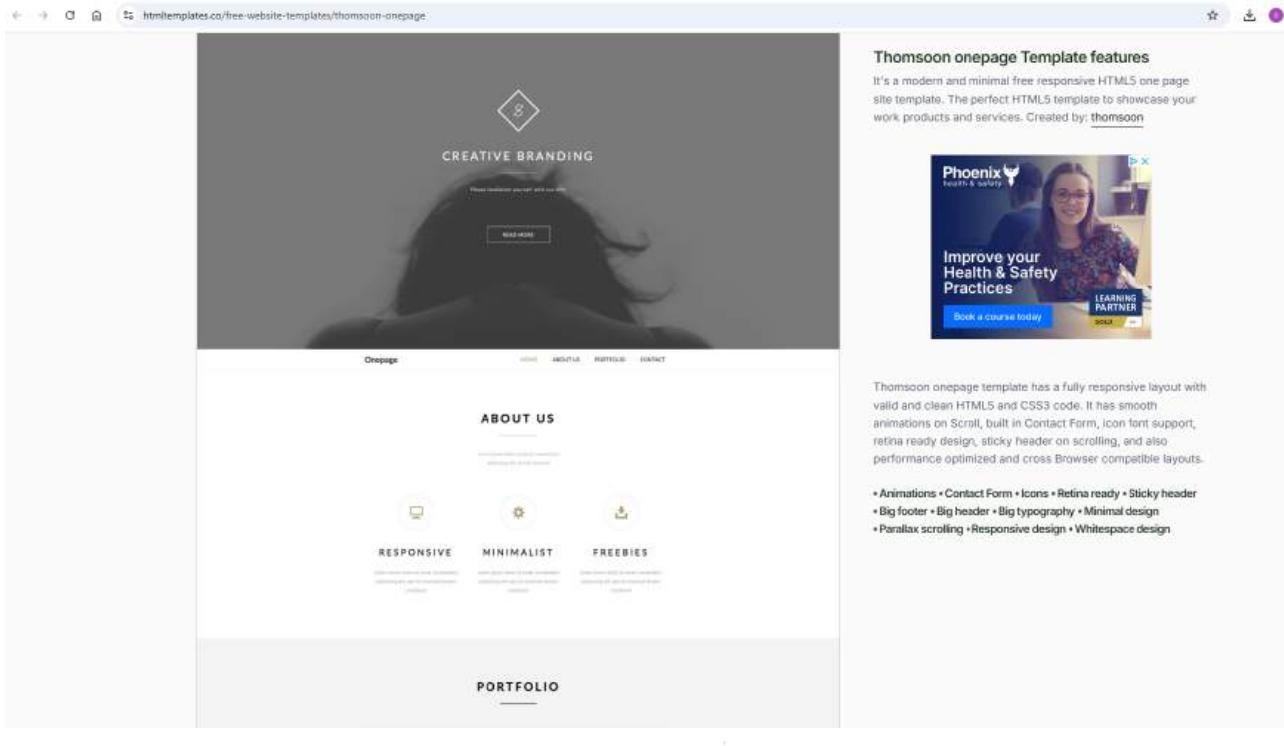
<https://html5up.net/>

<https://onepagelove.com/templates/free-templates>

<https://htmlrev.com/free-html-templates.html>

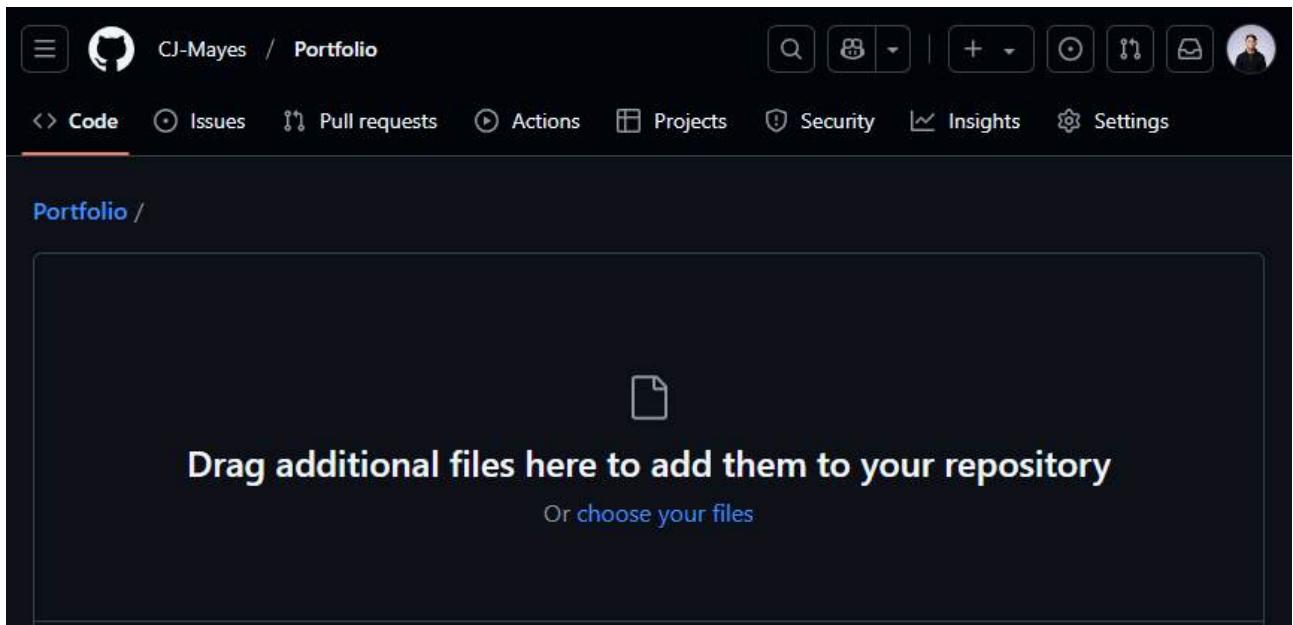
It will comprise generally of:

a css folder, images folder, js and most importantly an index.html file.  
If you want to follow along with the template I used it was [found here](#).



	File folder
	File folder
	File folder
	Chrome HTML Document
	Chrome HTML Document
	PHP File

You can add these files directly into the repo, by dragging them from the folder and onto the page.



Personally I think its a good idea to open an IDE of your choosing and login then clone these files into your repository instead.

```
Cloning into 'Portfolio'...
warning: You appear to have cloned an empty repository.
PS C:\Users\Cj.Mayes\PycharmProjects\Portfolio> 
```

Push all these files into the repository.  
That will be your template created.

```
To https://github.com/[REDACTED]/Portfolio.git
* [new branch]      main -> main
```

Next thing we want to turn the page on.

 Portfolio Public

Pin Unwatch 1

main 1 Branch 0 Tags Go to file Add file Code

CJ-Mayes template	95b0cbd · 2 minutes ago	1 Commit
css	template	2 minutes ago
img	template	2 minutes ago
js	template	2 minutes ago
api.html	template	2 minutes ago
index.html	template	2 minutes ago
mail.php	template	2 minutes ago

README

Add a README

Help people interested in this repository understand your project by adding a README.

Add a README

Go to settings and Github pages.  
Switch it to Github actions.

Code Issues Pull requests Actions Projects Security Insights Settings

Type / to search

GitHub Pages source saved.

General GitHub Pages

Access GitHub Pages is designed to host your personal, organization, or project pages from a GitHub repository.

Collaborators Moderation options Build and deployment

Code and automation Source GitHub Actions

Rules Use a suggested workflow, [browse all workflows](#), or [create your own](#).

Actions Webhooks Environments Codenames Pages

Next.js By GitHub Actions Package a Next.js site. Configure

NextJS By GitHub Actions Package a NextJS site. Configure

For the time being we will look to use a static HTML template.

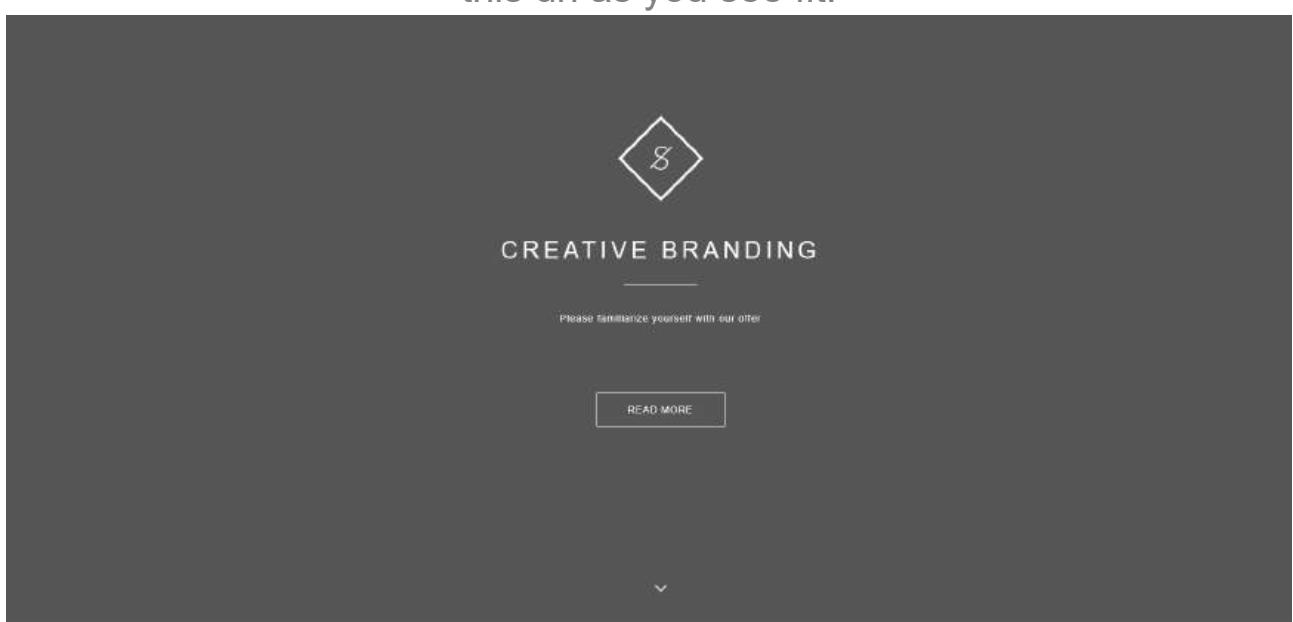
The screenshot shows a GitHub repository interface. At the top, there's a navigation bar with 'main' selected. Below it, a workflow named 'Create static.yml' is listed under 'Portfolios / github / workflows'. The workflow details page shows a single step: 'Create static.yml'. There are buttons for 'Go to file', 'Add file', and other repository actions.

The YML file will trigger the github actions to load the HTML page template we created.

The screenshot shows the GitHub Pages settings for a repository. On the left, a sidebar lists 'General', 'Access', 'Collaborators', 'Moderation options', 'Code and automation' (with 'Actions' selected), 'Environments', 'Codepaces', and 'Pages' (selected). The main area is titled 'GitHub Pages' and contains sections for 'Build and deployment' (using 'GitHub Actions'), 'Custom domain' (disabled), and 'Visibility' (set to 'Github Enterprise'). A note at the bottom says 'Your site was last deployed to the [github-pages](#) environment by the [Deploy static content to Pages](#) workflow.'

With that file created we can go back to pages and see the site is live.  
Click your custom link.

For now it will be your github name / Portfolio. Of course you can amend this url as you see fit.



So that is us set up!

But now we need to know how to

1. Update elements and items in the template itself now its live.

2. Know how to test locally before publishing live.

A static HTML site typically includes an index.html file, along with img, css, and js folders. Here's what each of these elements does:

The index.html file serves as the main entry point for your website. It contains the structure and content of your homepage, using HTML tags to define headings, paragraphs, links, and other elements. When someone visits your site's root URL, this is usually the first page they see.

The img folder is where you store all the images used on your website.

This includes photographs, logos, icons, and any other graphics.

Keeping images in a separate folder helps organize your site's assets and makes it easier to reference them in your HTML and CSS files.

The css folder contains your Cascading Style Sheets. These files define the visual appearance of your HTML elements, including colors, fonts, layouts, and responsive design rules. Separating your CSS from your HTML helps maintain a clean structure and makes it easier to update your site's design.

The js folder houses your JavaScript files. These add interactivity and dynamic behavior to your otherwise static HTML pages. JavaScript can be used for form validation, creating animations, updating content without reloading the page, and much more.

This structure helps keep your site organized and makes it easier to maintain and update over time. It's a common and effective way to set up a static HTML site.

Lets take a look at the simplist example for now.

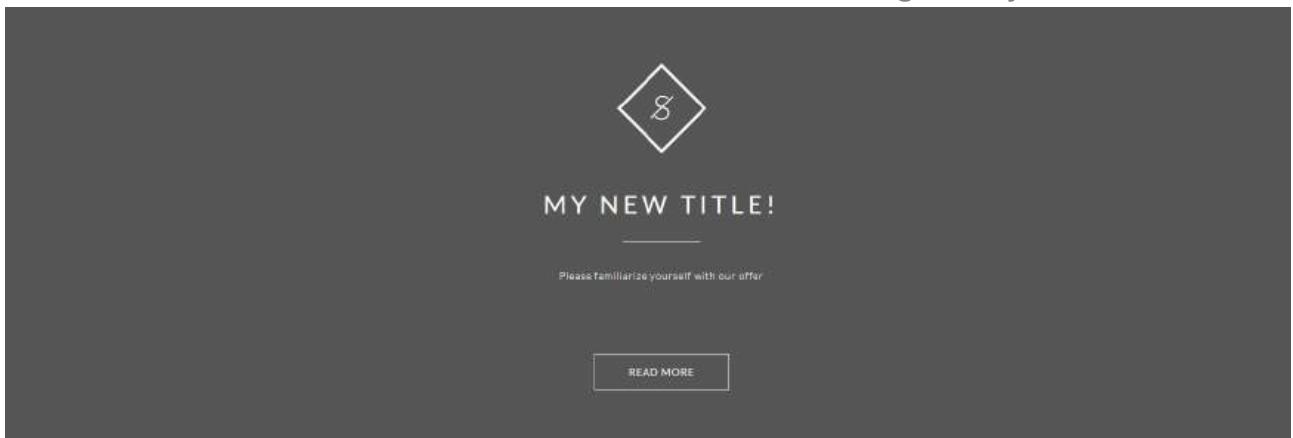
In the Index file you'll see a header. In the case of this template it reads "Creative Branding"

```
index.html
Portfolio
  38
  39      <!-- section start-page -->
  40
  41      <section class="start-page parallax-background" id="home">
  42
  43          <div class="opacity"></div> <!-- Opacity color -->
  44          <div class="content">
  45              <div class="text">
  46
  47                  <div class="logo"></div>
  48
  49                  <h1>creative branding</h1><hr/>
  50                  <p>Please familiarize yourself with our offer</p>
  51
  52                  <a href="#about-us"><div class="read-more">Read more</div></a>
  53
  54          </div>
  55          <div class="arrow-down"></div>
  56      </div>
  57
  58  </section>
  59
  60  <!-- section menu mobile -->
  61
  62  <section class="menu-media">
  63
  64      <div class="menu-content">
  65
  66          <div class="logo">OnePage</div>
  67
  68          <div class="icon"><a href="#"></a></div>
  69
  70      </div>
  71
  72
  73
  74
  75
  76
  77
  78
  79
  80
  81
  82
  83
  84
  85
  86
  87
  88
  89
  90
  91
  92
  93
  94
  95
  96
  97
  98
  99
  100
  101
  102
  103
  104
  105
  106
  107
  108
  109
  110
  111
  112
  113
  114
  115
  116
  117
  118
  119
  120
  121
  122
  123
  124
  125
  126
  127
  128
  129
  130
  131
  132
  133
  134
  135
  136
  137
  138
  139
  140
  141
  142
  143
  144
  145
  146
  147
  148
  149
  150
  151
  152
  153
  154
  155
  156
  157
  158
  159
  160
  161
  162
  163
  164
  165
  166
  167
  168
  169
  170
  171
  172
  173
  174
  175
  176
  177
  178
  179
  180
  181
  182
  183
  184
  185
  186
  187
  188
  189
  190
  191
  192
  193
  194
  195
  196
  197
  198
  199
  200
  201
  202
  203
  204
  205
  206
  207
  208
  209
  210
  211
  212
  213
  214
  215
  216
  217
  218
  219
  220
  221
  222
  223
  224
  225
  226
  227
  228
  229
  230
  231
  232
  233
  234
  235
  236
  237
  238
  239
  240
  241
  242
  243
  244
  245
  246
  247
  248
  249
  250
  251
  252
  253
  254
  255
  256
  257
  258
  259
  260
  261
  262
  263
  264
  265
  266
  267
  268
  269
  270
  271
  272
  273
  274
  275
  276
  277
  278
  279
  280
  281
  282
  283
  284
  285
  286
  287
  288
  289
  290
  291
  292
  293
  294
  295
  296
  297
  298
  299
  300
  301
  302
  303
  304
  305
  306
  307
  308
  309
  310
  311
  312
  313
  314
  315
  316
  317
  318
  319
  320
  321
  322
  323
  324
  325
  326
  327
  328
  329
  330
  331
  332
  333
  334
  335
  336
  337
  338
  339
  340
  341
  342
  343
  344
  345
  346
  347
  348
  349
  350
  351
  352
  353
  354
  355
  356
  357
  358
  359
  360
  361
  362
  363
  364
  365
  366
  367
  368
  369
  370
  371
  372
  373
  374
  375
  376
  377
  378
  379
  380
  381
  382
  383
  384
  385
  386
  387
  388
  389
  390
  391
  392
  393
  394
  395
  396
  397
  398
  399
  400
  401
  402
  403
  404
  405
  406
  407
  408
  409
  410
  411
  412
  413
  414
  415
  416
  417
  418
  419
  420
  421
  422
  423
  424
  425
  426
  427
  428
  429
  430
  431
  432
  433
  434
  435
  436
  437
  438
  439
  440
  441
  442
  443
  444
  445
  446
  447
  448
  449
  450
  451
  452
  453
  454
  455
  456
  457
  458
  459
  460
  461
  462
  463
  464
  465
  466
  467
  468
  469
  470
  471
  472
  473
  474
  475
  476
  477
  478
  479
  480
  481
  482
  483
  484
  485
  486
  487
  488
  489
  490
  491
  492
  493
  494
  495
  496
  497
  498
  499
  500
  501
  502
  503
  504
  505
  506
  507
  508
  509
  510
  511
  512
  513
  514
  515
  516
  517
  518
  519
  520
  521
  522
  523
  524
  525
  526
  527
  528
  529
  530
  531
  532
  533
  534
  535
  536
  537
  538
  539
  540
  541
  542
  543
  544
  545
  546
  547
  548
  549
  550
  551
  552
  553
  554
  555
  556
  557
  558
  559
  560
  561
  562
  563
  564
  565
  566
  567
  568
  569
  570
  571
  572
  573
  574
  575
  576
  577
  578
  579
  580
  581
  582
  583
  584
  585
  586
  587
  588
  589
  590
  591
  592
  593
  594
  595
  596
  597
  598
  599
  600
  601
  602
  603
  604
  605
  606
  607
  608
  609
  610
  611
  612
  613
  614
  615
  616
  617
  618
  619
  620
  621
  622
  623
  624
  625
  626
  627
  628
  629
  630
  631
  632
  633
  634
  635
  636
  637
  638
  639
  640
  641
  642
  643
  644
  645
  646
  647
  648
  649
  650
  651
  652
  653
  654
  655
  656
  657
  658
  659
  660
  661
  662
  663
  664
  665
  666
  667
  668
  669
  670
  671
  672
  673
  674
  675
  676
  677
  678
  679
  680
  681
  682
  683
  684
  685
  686
  687
  688
  689
  690
  691
  692
  693
  694
  695
  696
  697
  698
  699
  700
  701
  702
  703
  704
  705
  706
  707
  708
  709
  710
  711
  712
  713
  714
  715
  716
  717
  718
  719
  720
  721
  722
  723
  724
  725
  726
  727
  728
  729
  730
  731
  732
  733
  734
  735
  736
  737
  738
  739
  740
  741
  742
  743
  744
  745
  746
  747
  748
  749
  750
  751
  752
  753
  754
  755
  756
  757
  758
  759
  760
  761
  762
  763
  764
  765
  766
  767
  768
  769
  770
  771
  772
  773
  774
  775
  776
  777
  778
  779
  780
  781
  782
  783
  784
  785
  786
  787
  788
  789
  790
  791
  792
  793
  794
  795
  796
  797
  798
  799
  800
  801
  802
  803
  804
  805
  806
  807
  808
  809
  810
  811
  812
  813
  814
  815
  816
  817
  818
  819
  820
  821
  822
  823
  824
  825
  826
  827
  828
  829
  830
  831
  832
  833
  834
  835
  836
  837
  838
  839
  840
  841
  842
  843
  844
  845
  846
  847
  848
  849
  850
  851
  852
  853
  854
  855
  856
  857
  858
  859
  860
  861
  862
  863
  864
  865
  866
  867
  868
  869
  870
  871
  872
  873
  874
  875
  876
  877
  878
  879
  880
  881
  882
  883
  884
  885
  886
  887
  888
  889
  890
  891
  892
  893
  894
  895
  896
  897
  898
  899
  900
  901
  902
  903
  904
  905
  906
  907
  908
  909
  910
  911
  912
  913
  914
  915
  916
  917
  918
  919
  920
  921
  922
  923
  924
  925
  926
  927
  928
  929
  930
  931
  932
  933
  934
  935
  936
  937
  938
  939
  940
  941
  942
  943
  944
  945
  946
  947
  948
  949
  950
  951
  952
  953
  954
  955
  956
  957
  958
  959
  960
  961
  962
  963
  964
  965
  966
  967
  968
  969
  970
  971
  972
  973
  974
  975
  976
  977
  978
  979
  980
  981
  982
  983
  984
  985
  986
  987
  988
  989
  990
  991
  992
  993
  994
  995
  996
  997
  998
  999
  1000
  1001
  1002
  1003
  1004
  1005
  1006
  1007
  1008
  1009
  1010
  1011
  1012
  1013
  1014
  1015
  1016
  1017
  1018
  1019
  1020
  1021
  1022
  1023
  1024
  1025
  1026
  1027
  1028
  1029
  1030
  1031
  1032
  1033
  1034
  1035
  1036
  1037
  1038
  1039
  1040
  1041
  1042
  1043
  1044
  1045
  1046
  1047
  1048
  1049
  1050
  1051
  1052
  1053
  1054
  1055
  1056
  1057
  1058
  1059
  1060
  1061
  1062
  1063
  1064
  1065
  1066
  1067
  1068
  1069
  1070
  1071
  1072
  1073
  1074
  1075
  1076
  1077
  1078
  1079
  1080
  1081
  1082
  1083
  1084
  1085
  1086
  1087
  1088
  1089
  1090
  1091
  1092
  1093
  1094
  1095
  1096
  1097
  1098
  1099
  1100
  1101
  1102
  1103
  1104
  1105
  1106
  1107
  1108
  1109
  1110
  1111
  1112
  1113
  1114
  1115
  1116
  1117
  1118
  1119
  1120
  1121
  1122
  1123
  1124
  1125
  1126
  1127
  1128
  1129
  1130
  1131
  1132
  1133
  1134
  1135
  1136
  1137
  1138
  1139
  1140
  1141
  1142
  1143
  1144
  1145
  1146
  1147
  1148
  1149
  1150
  1151
  1152
  1153
  1154
  1155
  1156
  1157
  1158
  1159
  1160
  1161
  1162
  1163
  1164
  1165
  1166
  1167
  1168
  1169
  1170
  1171
  1172
  1173
  1174
  1175
  1176
  1177
  1178
  1179
  1180
  1181
  1182
  1183
  1184
  1185
  1186
  1187
  1188
  1189
  1190
  1191
  1192
  1193
  1194
  1195
  1196
  1197
  1198
  1199
  1200
  1201
  1202
  1203
  1204
  1205
  1206
  1207
  1208
  1209
  1210
  1211
  1212
  1213
  1214
  1215
  1216
  1217
  1218
  1219
  1220
  1221
  1222
  1223
  1224
  1225
  1226
  1227
  1228
  1229
  1230
  1231
  1232
  1233
  1234
  1235
  1236
  1237
  1238
  1239
  1240
  1241
  1242
  1243
  1244
  1245
  1246
  1247
  1248
  1249
  1250
  1251
  1252
  1253
  1254
  1255
  1256
  1257
  1258
  1259
  1260
  1261
  1262
  1263
  1264
  1265
  1266
  1267
  1268
  1269
  1270
  1271
  1272
  1273
  1274
  1275
  1276
  1277
  1278
  1279
  1280
  1281
  1282
  1283
  1284
  1285
  1286
  1287
  1288
  1289
  1290
  1291
  1292
  1293
  1294
  1295
  1296
  1297
  1298
  1299
  1300
  1301
  1302
  1303
  1304
  1305
  1306
  1307
  1308
  1309
  1310
  1311
  1312
  1313
  1314
  1315
  1316
  1317
  1318
  1319
  1320
  1321
  1322
  1323
  1324
  1325
  1326
  1327
  1328
  1329
  1330
  1331
  1332
  1333
  1334
  1335
  1336
  1337
  1338
  1339
  1340
  1341
  1342
  1343
  1344
  1345
  1346
  1347
  1348
  1349
  1350
  1351
  1352
  1353
  1354
  1355
  1356
  1357
  1358
  1359
  1360
  1361
  1362
  1363
  1364
  1365
  1366
  1367
  1368
  1369
  1370
  1371
  1372
  1373
  1374
  1375
  1376
  1377
  1378
  1379
  1380
  1381
  1382
  1383
  1384
  1385
  1386
  1387
  1388
  1389
  1390
  1391
  1392
  1393
  1394
  1395
  1396
  1397
  1398
  1399
  1400
  1401
  1402
  1403
  1404
  1405
  1406
  1407
  1408
  1409
  1410
  1411
  1412
  1413
  1414
  1415
  1416
  1417
  1418
  1419
  1420
  1421
  1422
  1423
  1424
  1425
  1426
  1427
  1428
  1429
  1430
  1431
  1432
  1433
  1434
  1435
  1436
  1437
  1438
  1439
  1440
  1441
  1442
  1443
  1444
  1445
  1446
  1447
  1448
  1449
  1450
  1451
  1452
  1453
  1454
  1455
  1456
  1457
  1458
  1459
  1460
  1461
  1462
  1463
  1464
  1465
  1466
  1467
  1468
  1469
  1470
  1471
  1472
  1473
  1474
  1475
  1476
  1477
  1478
  1479
  1480
  1481
  1482
  1483
  1484
  1485
  1486
  1487
  1488
  1489
  1490
  1491
  1492
  1493
  1494
  1495
  1496
  1497
  1498
  1499
  1500
  1501
  1502
  1503
  1504
  1505
  1506
  1507
  1508
  1509
  1510
  1511
  1512
  1513
  1514
  1515
  1516
  1517
  1518
  1519
  1520
  1521
  1522
  1523
  1524
  1525
  1526
  1527
  1528
  1529
  1530
  1531
  1532
  1533
  1534
  1535
  1536
  1537
  1538
  1539
  1540
  1541
  1542
  1543
  1544
  1545
  1546
  1547
  1548
  1549
  1550
  1551
  1552
  1553
  1554
  1555
  1556
  1557
  1558
  1559
  1560
  1561
  1562
  1563
  1564
  1565
  1566
  1567
  1568
  1569
  1570
  1571
  1572
  1573
  1574
  1575
  1576
  1577
  1578
  1579
  1580
  1581
  1582
  1583
  1584
  1585
  1586
  1587
  1588
  1589
  1590
  1591
  1592
  1593
  1594
  1595
  1596
  1597
  1598
  1599
  1600
  1601
  1602
  1603
  1604
  1605
  1606
  1607
  1608
  1609
  1610
  1611
  1612
  1613
  1614
  1615
  1616
  1617
  1618
  1619
  1620
  1621
  1622
  1623
  1624
  1625
  1626
  1627
  1628
  1629
  1630
  1631
  1632
  1633
  1634
  1635
  1636
  1637
  1638
  1639
  1640
  1641
  1642
  1643
  1644
  1645
  1646
  1647
  1648
  1649
  1650
  1651
  1652
  1653
  1654
  1655
  1656
  1657
  1658
  1659
  1660
  1661
  1662
  1663
  1664
  1665
  1666
  1667
  1668
  1669
  1670
  1671
  1672
  1673
  1674
  1675
  1676
  1677
  1678
  1679
  1680
  1681
  1682
  1683
  1684
  1685
  1686
  1687
  1688
  1689
  1690
  1691
  1692
  1693
  1694
  1695
  1696
  1697
  1698
  1699
  1700
  1701
  1702
  1703
  1704
  1705
  1706
  1707
  1708
  1709
  1710
  1711
  1712
  1713
  1714
  1715
  1716
  1717
  1718
  1719
  1720
  1721
  1722
  1723
  1724
  1725
  1726
  1727
  1728
  1729
  1730
  1731
  1732
  1733
```

```
File: 123 index.html

29      
30
31      <section class="start-page parallax-background" id="home">
32
33          <div class="opacity"></div> <!-- Opacity color -->
34          <div class="content">
35              <div class="text">
36
37                  <div class="logo"></div>
38
39                  <h1>My New Title!</h1><hr/>
40                  <p>Please familiarize yourself with our offer</p>
41
42                  <a href="#about-us"><div class="read-more">Read more</div></a>
43
44              </div>
45              <div class="arrow-down"></div>
46          </div>
47
48      </section>
49
50      <!-- section menu/mobile --&gt;
51
52      &lt;section class="menu-media"&gt;
53
54          &lt;div class="menu-content"&gt;
55
56              &lt;div class="Logo"&gt;OnePage&lt;/div&gt;
57
58              &lt;div class="icon"&gt;&lt;a href="#"&gt;&lt;img src="img/icons/menu-media.png"/&gt;&lt;/a&gt;&lt;/div&gt;
59
60          &lt;/div&gt;
61
62      &lt;/section&gt;</pre>
```

Now that's all well and good as we can see the changes.... but this would take us forever and a million commits to get any work done.

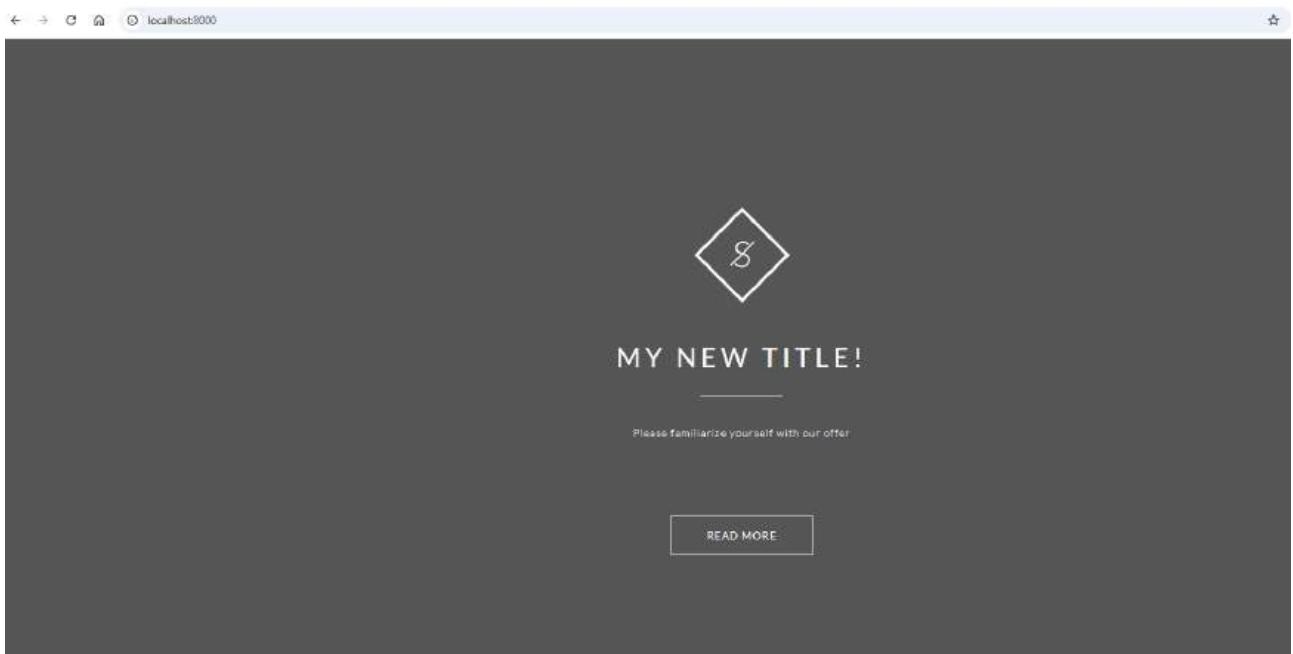


So let's actually test this locally before we publish it up to our repo.

```
PS C:\Users\Cj.Mayes\PycharmProjects\Portfolio\Portfolio> python -m http.server
Serving HTTP on :: port 8000 (http://[::]:8000/) ...
```

```
python -m http.server
```

Start the local server in your terminal.



Now you can amend your files as you see fit, before you push them live to your repository.

To push your edited files live generally follow the below commands.  
git status: This command shows you the current state of your working directory and staging area. It tells you which files have been modified, which are staged for commit, and which are untracked. It's a good way to get an overview of your changes before committing.

#### Git status

git add: This command stages changes for commit. You can add specific files (git add filename) or all changes (git add .). Staging allows you to choose which changes you want to include in your next commit. It's like preparing a package before sending it.

#### Git add .

git commit: This creates a new commit with the changes you've staged. A commit is like a snapshot of your project at a specific point in time. You typically include a message describing the changes: git commit -m "Your commit message here". This helps you and others understand what changes were made and why.

#### Git commit -m "adding my new title to the index.html"

git push: This command uploads your local commits to a remote repository, often on platforms like GitHub or GitLab. It's how you share

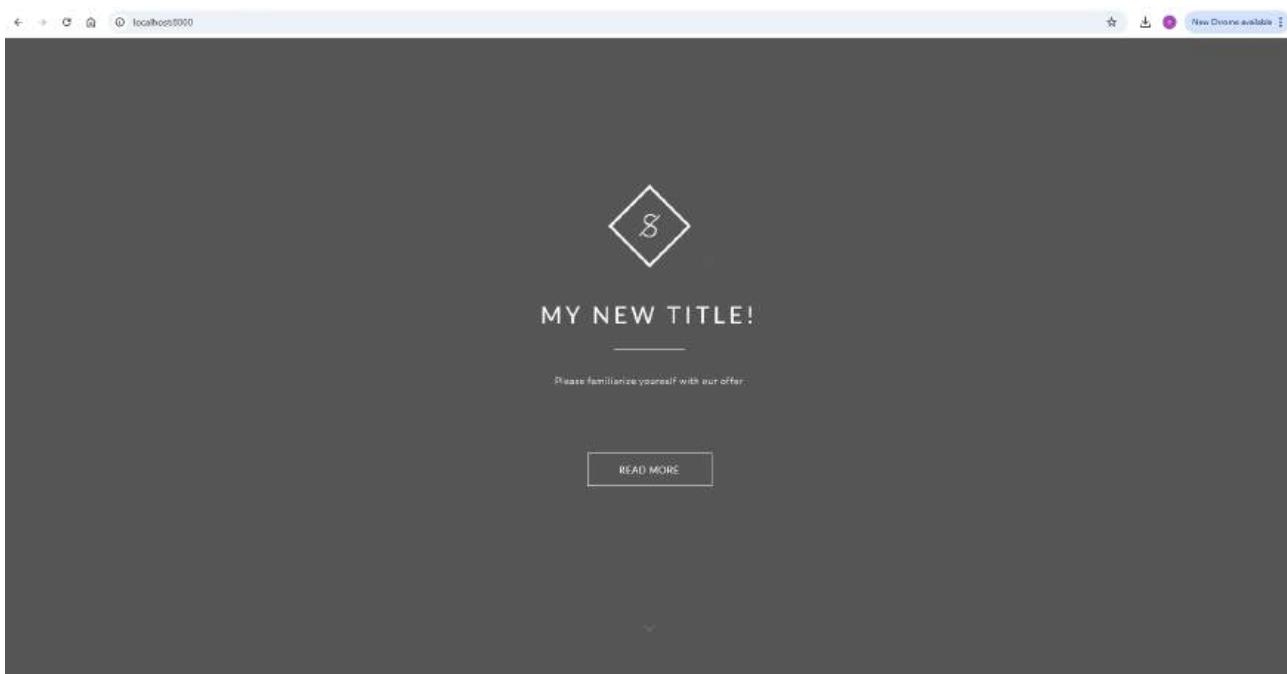
your changes with others or update the central version of your project.

The basic form is `git push origin branch-name`, where ‘origin’ is the default name for the remote repository and ‘branch-name’ is the branch you’re pushing.

#### Git push

These commands are often used together in a workflow:

1. Make changes to your files
2. Check status with `git status`
3. Stage changes with `git add`
4. Commit changes with `git commit`
5. Push changes to the remote with `git push`



Thats our whistlestop tour to creating a github 1 pager from a template.

Going further.

- Use a template of your choosing.
- Update some of the titles and headers.
- Amend the style of the page.
- Add in your images to the portfolio page.
- Create a custom url for the github page.

LOGGING OFF,

CJ

SVG PACKED BUBBLES IN TABLEAU

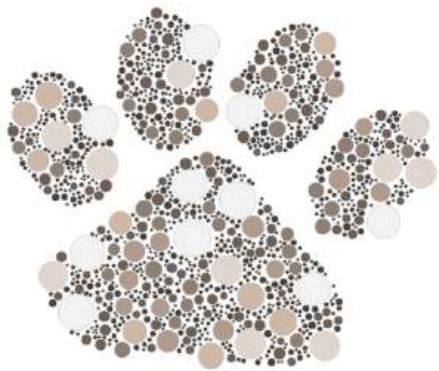
Hi all,

A fun data art style blog this week.  
This idea was heavily inspired from The Flerlage Twins btw. You can check out their blog [here](#).

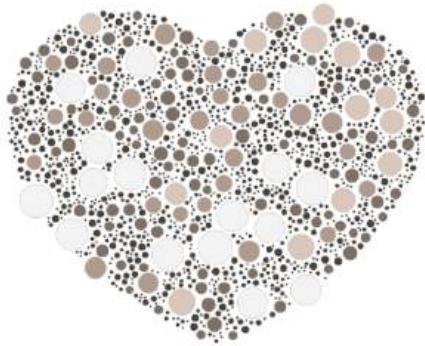
It was written 5 years ago now and still works perfectly fine, but i wanted to give my own shapes a go, and to see if I could get some code of my own going.

The way this blog will differ from the Twins is that we will look purely at Vector style SVG's. By that I mean filling a blank shape with packed circles.

SVG Packed Bubble Blog  
*Bengie's Paw Print | 1066 circles*



*Reasons To Love Bengal | 1604 circles*

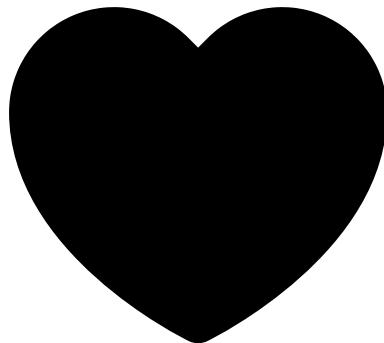


It won't touch on some of the more artistic flair that the twins had with recreating actual pieces of art and assigning each of the different hex codes.

For this to work, you will need.

- An SVG file.
- The code out of the repository (and the ability to run python code)
- Tableau (Optional)

Let's take this simple heart shape.



What we want to end up with is a bunch of randomly packed bubbles in this heart vector.



But then also I thought what if the vector is made up of multiple components, like this paw shape. So ended up revisiting the code to make sure that it takes the different "shapes" into account.

```
import random
import math
import os
from svgpath import svg2paths
import xml.etree.ElementTree as ET
import csv

def parse_svg(file_path):
    if not os.path.exists(file_path):
        raise FileNotFoundError(f"The file {file_path} does not exist.")

    paths, attributes = svg2paths(file_path)

    if not paths:
        raise ValueError("No path found in SVG file")

    boundaries = []
    for path in paths:
        boundary = []
        num_samples = max(100, int(path.length() / 10))
        boundary.extend([(path.point(t/num_samples)) for t in range(num_samples)])
        boundary = [(p.real, p.imag) for p in boundary]
        boundaries.append(boundary)

    # Scale all boundaries
    all_points = [p for b in boundaries for p in b]
    x_min = min(p[0] for p in all_points)
    x_max = max(p[0] for p in all_points)
    y_min = min(p[1] for p in all_points)
    y_max = max(p[1] for p in all_points)

    scale = 500 / max(x_max - x_min, y_max - y_min)
    scaled_boundaries = []
    for boundary in boundaries:
        scaled_boundary = [((p[0] - x_min) * scale, (p[1] - y_min) * scale) for p in boundary]
        scaled_boundaries.append(scaled_boundary)

    return scaled_boundaries

def is_inside_boundary(x, y, boundary):
    inside = False
    for i in range(len(boundary)):
        j = (i + 1) % len(boundary)
        if ((boundary[i][1] > y) != (boundary[j][1] > y)) and \
            ((x < (boundary[j][0] - boundary[i][0]) * (y - boundary[i][1])) / \
             (boundary[j][1] - boundary[i][1]) + boundary[i][0]):
            inside = not inside
    return inside

def pack_circles(boundary, min_radius=2, max_radius=25, max_attempts=10000):
    circles = []
    attempts = 0

    x_min = min(p[0] for p in boundary)
    x_max = max(p[0] for p in boundary)
    y_min = min(p[1] for p in boundary)
    y_max = max(p[1] for p in boundary)

    while attempts < max_attempts:
        radius = random.uniform(min_radius, max_radius)
        x = random.uniform(x_min, x_max)
        y = random.uniform(y_min, y_max)

        if is_inside_boundary(x, y, boundary):
            overlapping = False
            for circle in circles:
                distance = math.sqrt((x - circle[0])**2 + (y - circle[1])**2)
                if distance < radius + circle[2]:
                    overlapping = True
                    break

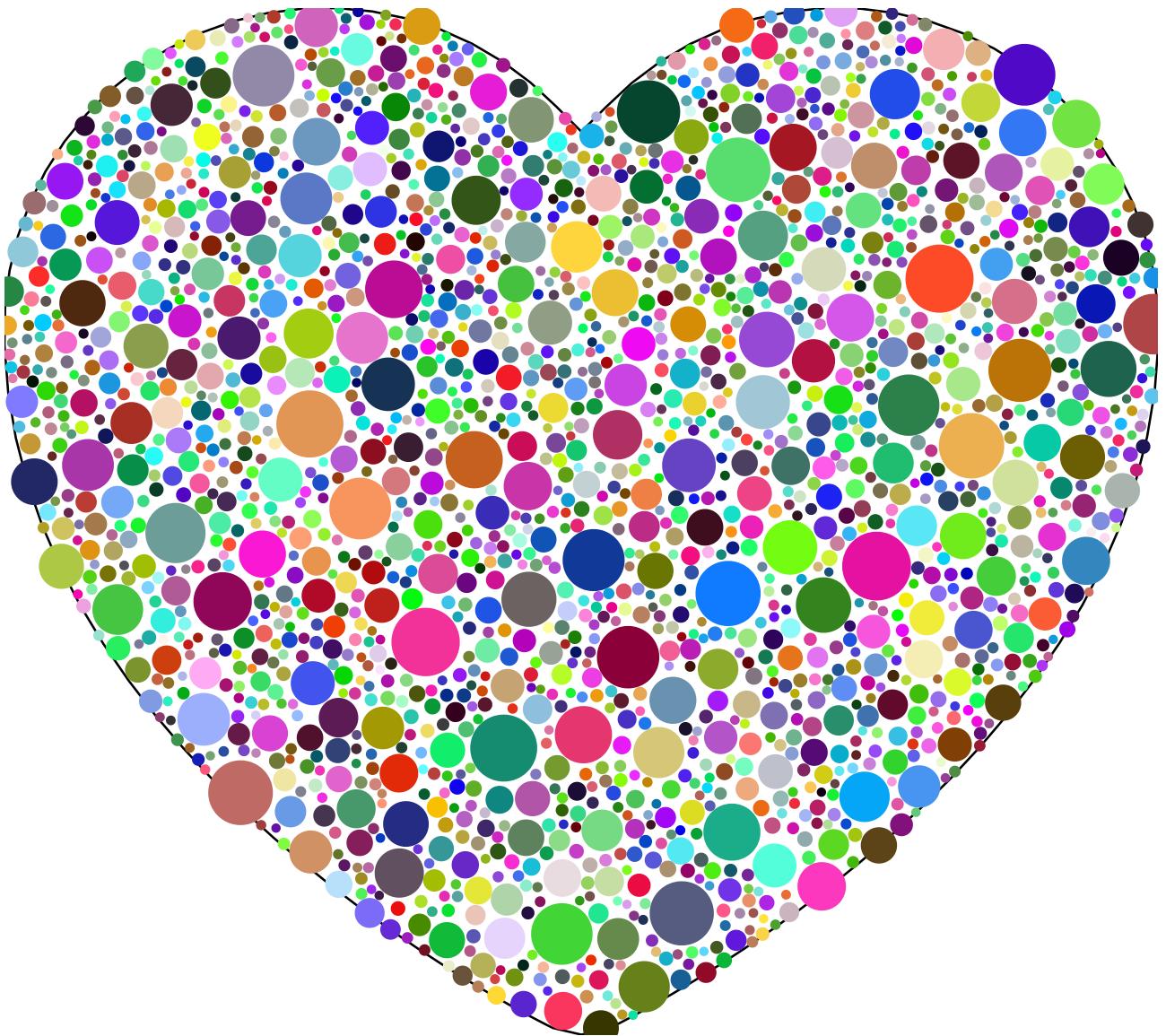
            if not overlapping:
                circles.append((x, y, radius))
                attempts = 0
            else:
                attempts += 1
        else:
```

The code in the repo creates a new SVG image with packed circles inside a boundary defined by an input SVG file. Here's a breakdown of its main components and functionality:

1. Imports: The code uses various libraries including random, math, os, svgpathutils, and xml.etree.ElementTree.
2. Try block: Handles exceptions. If a file is not found, it prints an error message and returns. If a ValueError is raised, it prints the error message and returns. If an Exception is raised, it prints an unexpected error occurred and returns.
3. If not boundaries: Checks if boundaries are provided. If not, it prints an error message and returns. Otherwise, it determines if a point (x, y) is inside a given boundary using the ray-casting algorithm.
4. For loop: Iterates over each boundary. Calls the pack\_circles function for each boundary.
5. All points: Ensures circles don't overlap and are fully within the boundary. (Note: You'll see that the circle boundary not necessarily the outer edge!)
6. For loop: Iterates over each boundary. Adds the original boundary paths and the new circles to the SVG.
7. For loop: Assigns random colors to the circles.
8. Error handling: Includes try-except blocks to handle potential errors like file not found or invalid SVG content.
9. Main execution: Prints the output file path and the number of circles created. Sets input and output file paths. Calls the create\_packed\_circle\_svg function to generate the new SVG.

This code essentially takes an SVG shape, fills it with randomly sized and positioned circles that don't overlap, and outputs a new SVG file with these packed circles inside the original shape.

```
# Updated input and output file paths
input_file = '/.../heart-solid.svg'
output_file = '/.../heart-solid-output.svg'
create_packed_circle_svg(input_file, output_file)
```



Above is an example output of the SVG..... (Excuse the randomised colour!) This heart contains 2033 circles!

we can tweak the size of the circles, through the min and max radius range. This heart only contains 269!

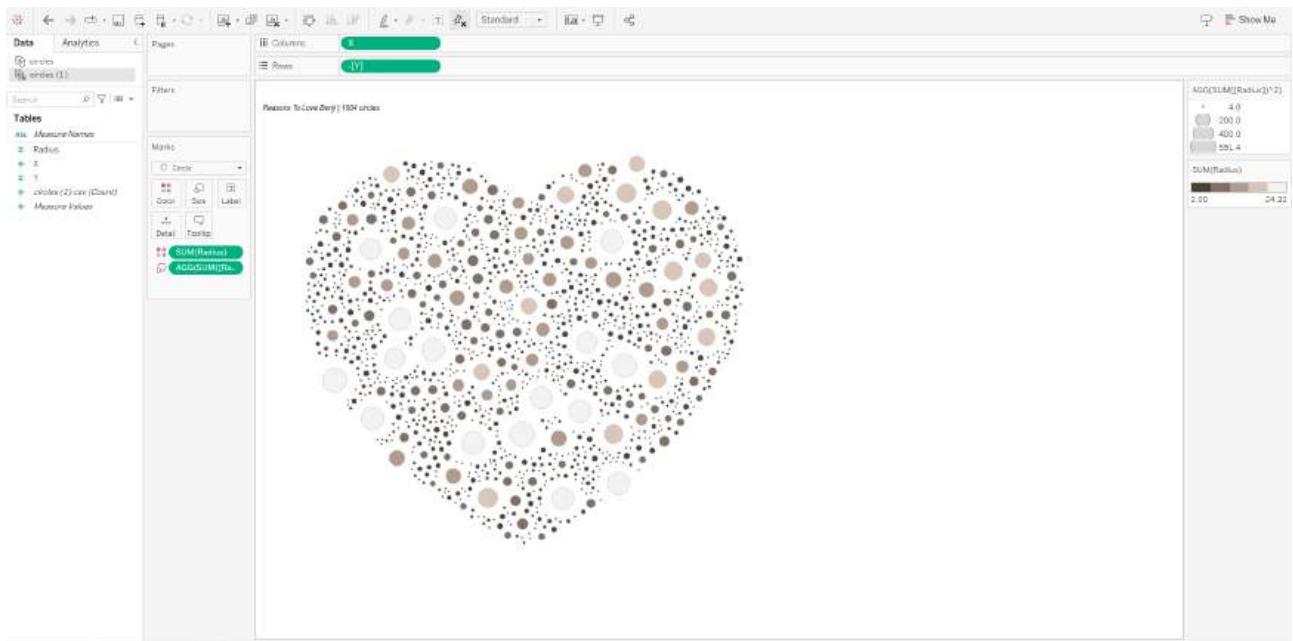


The code is just a matter of updating the input and output file.  
It will print out a statement telling you the SVG has saved, and the  
number of circles created as part of it.

It will then export the file required to re-build it in Tableau.

TABLEAU

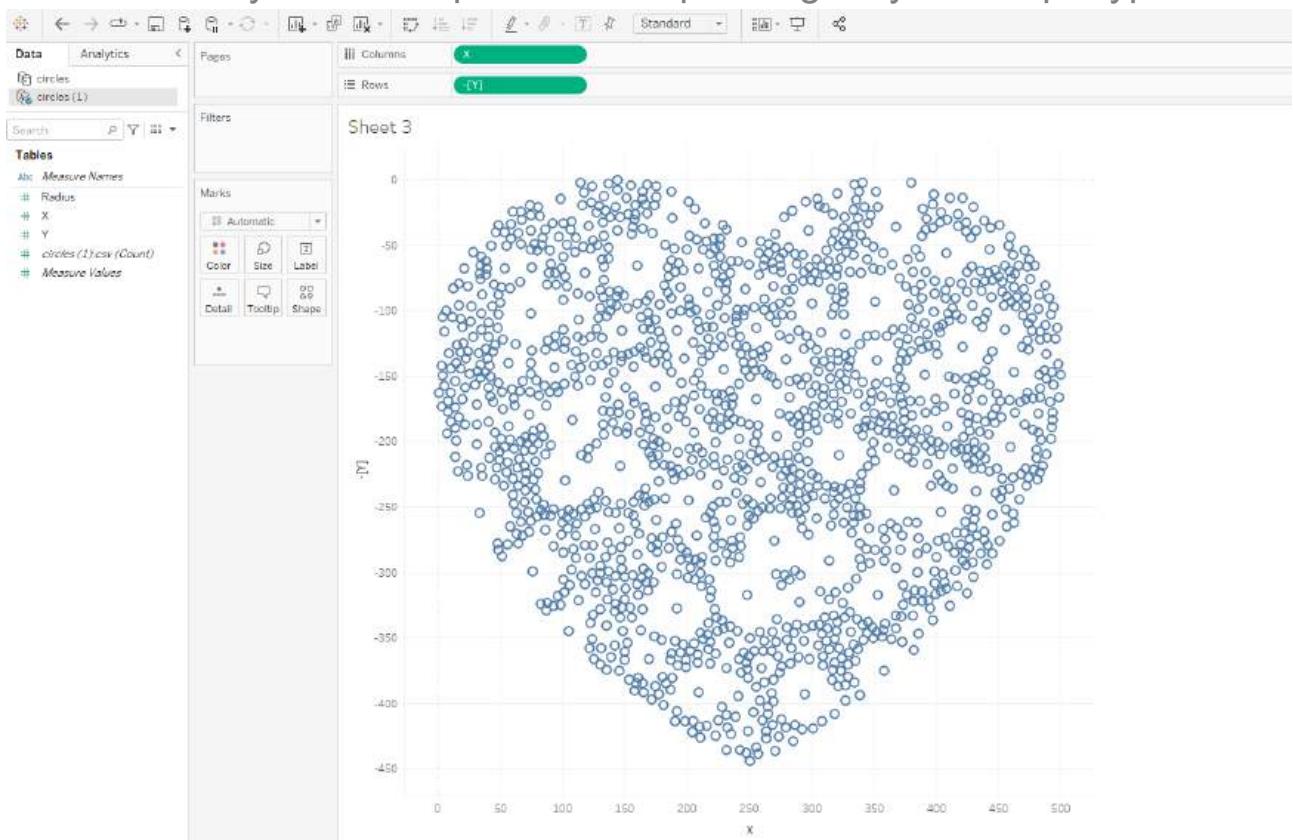
The easy bit!



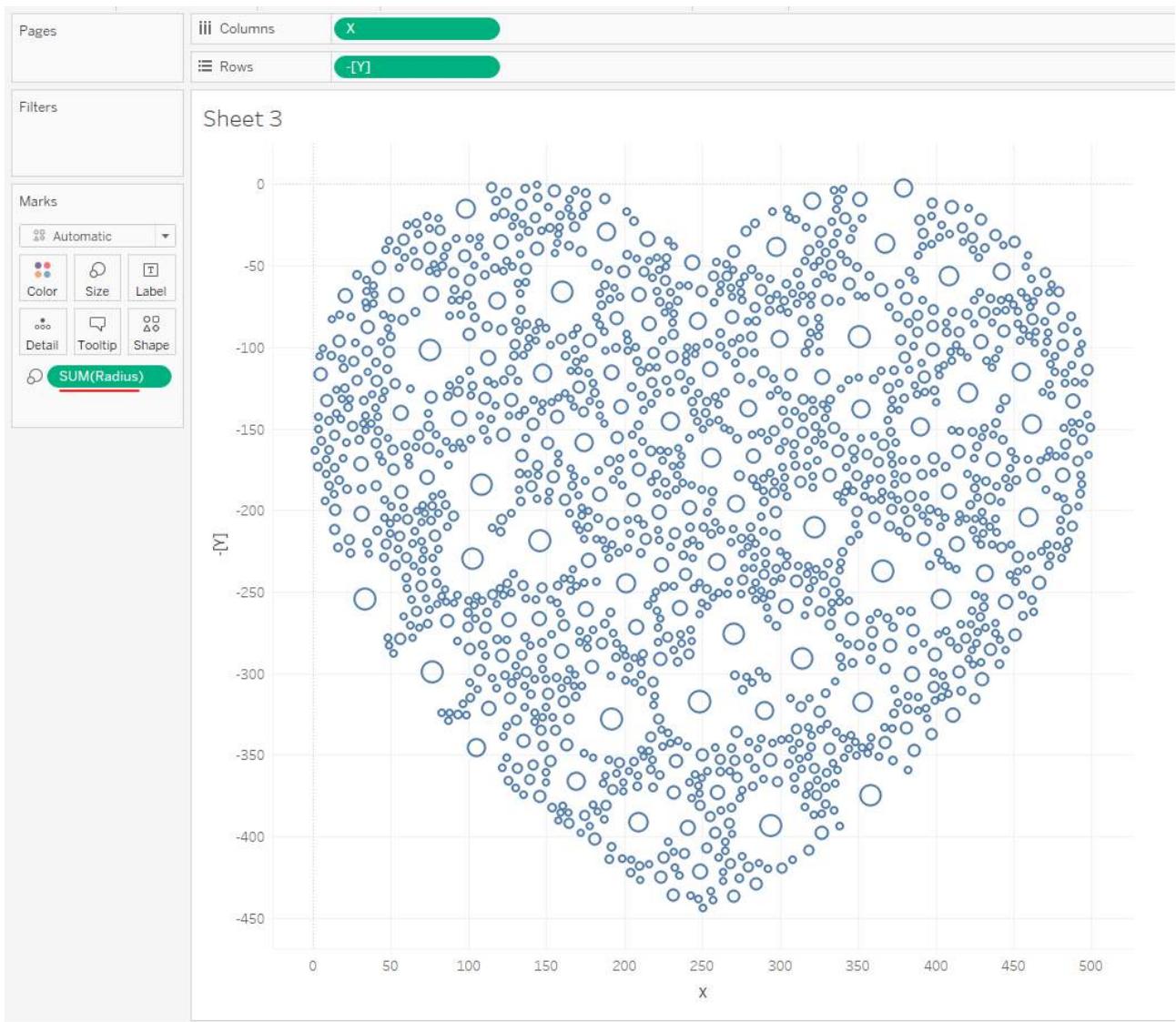
With your new exported file, connect to it in Tableau.

Put X onto Columns and Y onto Row.

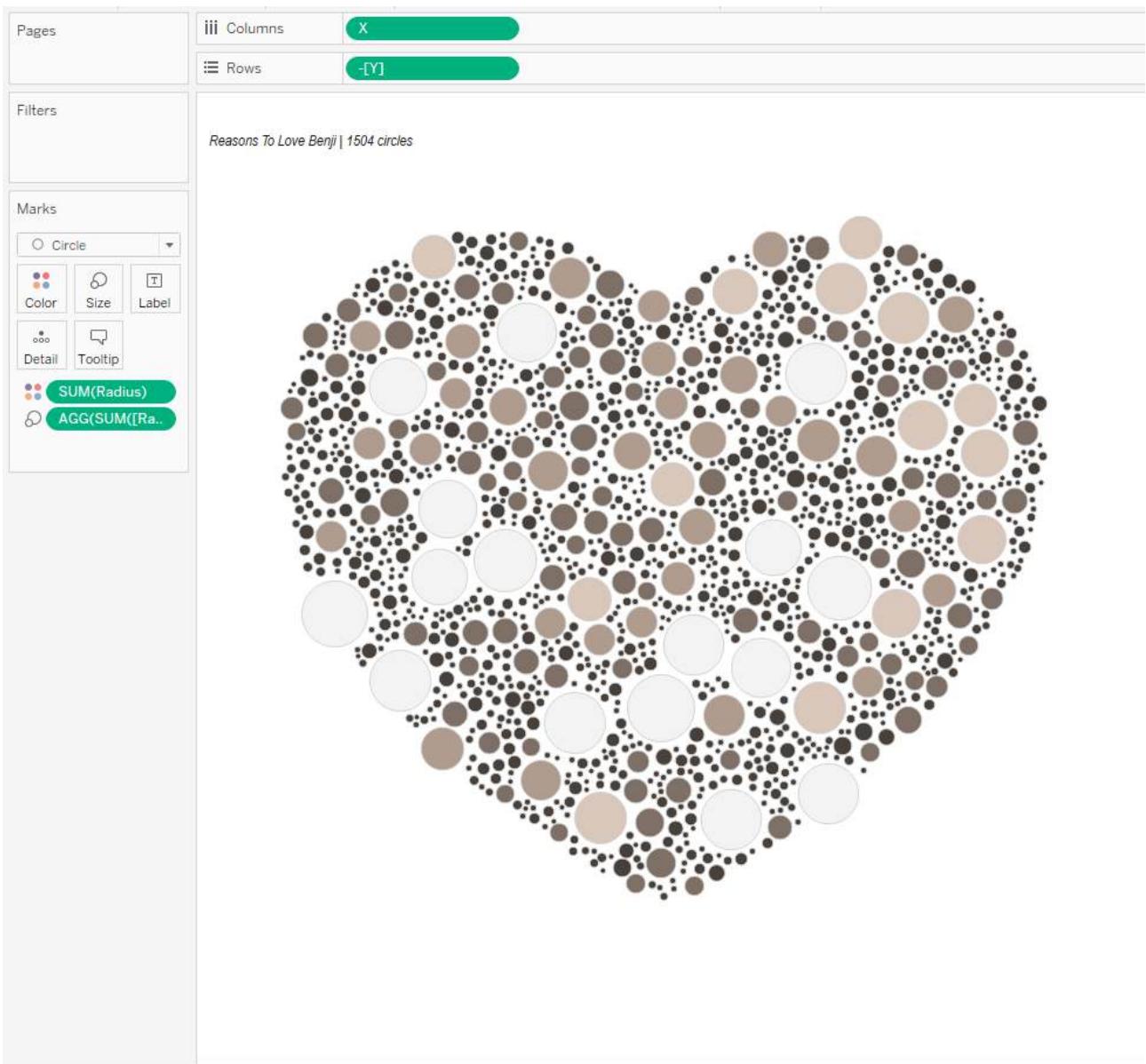
You may need to flip the axis depending on your shape type.



Notice how the axis goes between 0 and 500, well we set that also within the code. To tidy it up in the sheet it's probably worth fixing your axis. On the Tableau report example attached at the top of the page you'll see I fix both axis between -50 and 500.



Adding Radius to size hasn't quite filled the gaps as we probably would like, so I opted for doing Radius<sup>2</sup>.



& There we have it. A packed SVG shape in Tableau.  
Going further:

- (hard) Why not try and amend the code so that you can configure how many circles you ideally want and it add and removes circles to reach that number.
- (intermediate) Try your own SVG file
- (easy) Amend the code to print more or less circles using the min and max radius.

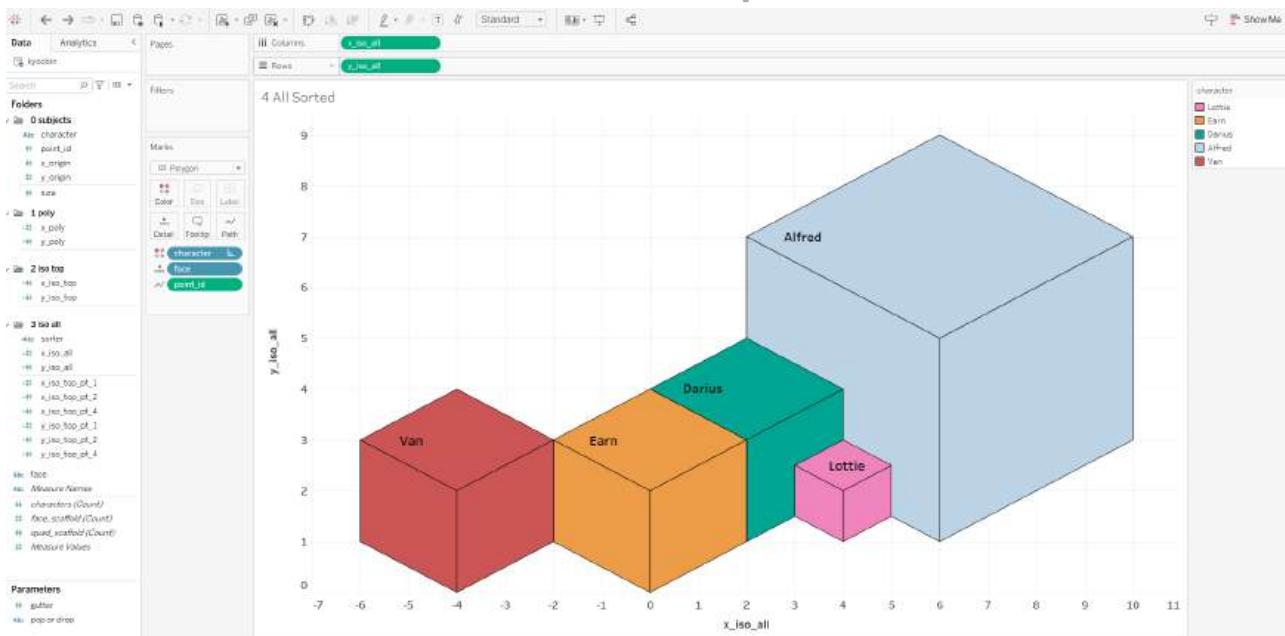
LOGGING OFF,  
CJ

## ISOMETRIC CHARTS IN TABLEAU WITH MATT ANTOLINE

Hi all,

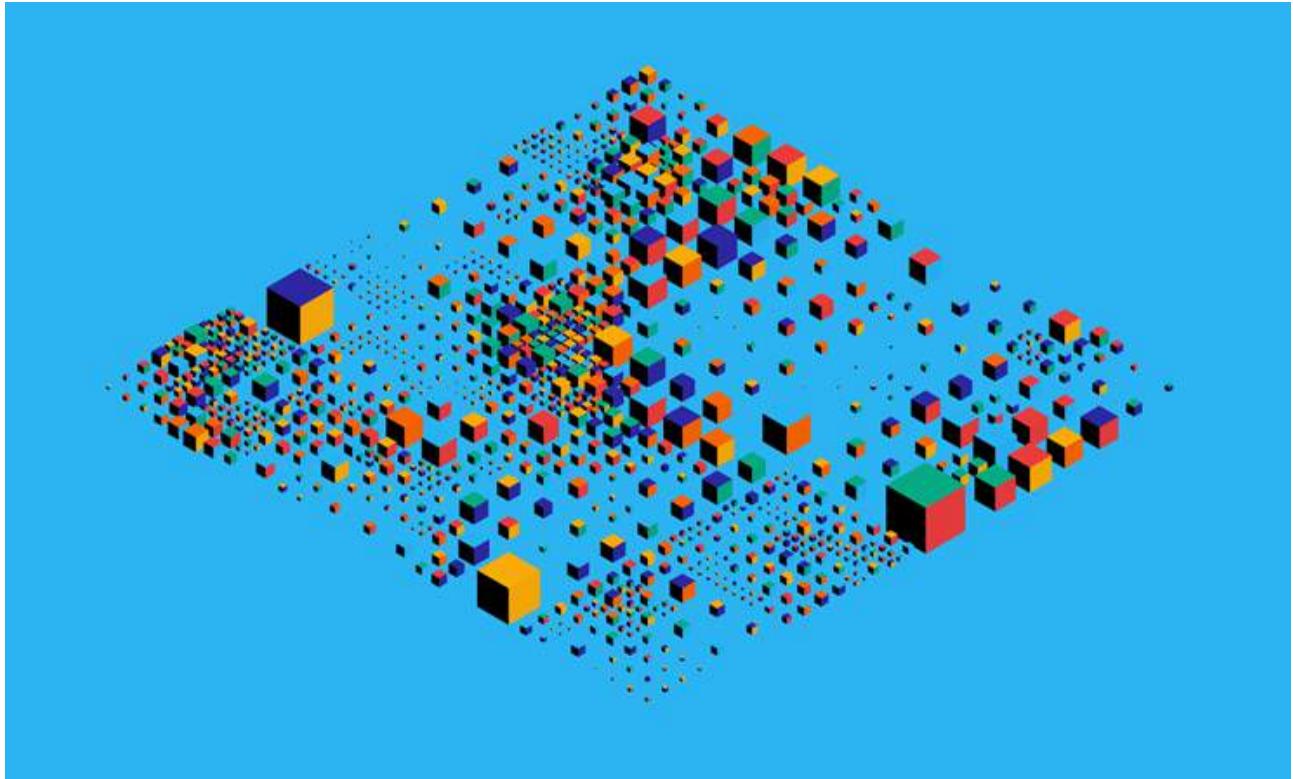
I don't say this lightly.... this might be one of my favourite technical guest blogs I've ever had the chance to host on the site.

I got to connect with Matt on a whim through sending him a message on Tableau Public, after Zach Bowders has tweeted about his work. Fast forward, Matt was kind enough to share his learning with everyone in blog form to learn how to create Isometric charts in tableau. Learn a little bit about Matt & then how to create your own Isometric visual & be sure to check out his profile [here](#).

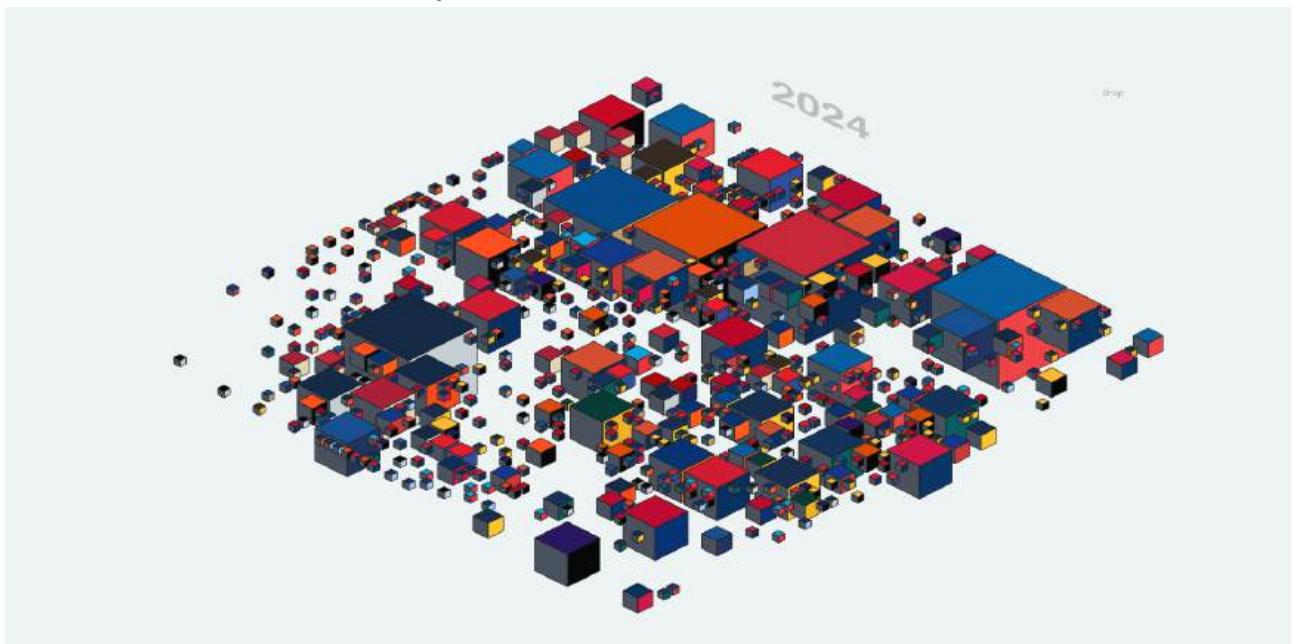


My day job is healthcare data. Hospital nerd stuff. When the nice old ladies at church ask me what I do for a living, I say, "My friends and I count sick people".... because the nice old ladies don't want to hear about entity relationship diagrams and data marts and feed ingestion and case statements and naming conventions and KPIs and outliers and have you tried clearing your cache, doctor?

By the end of a typical work day, I don't want to hear or think about those things either. I just want to make something pretty. Consequently my Tableau Public portfolio has largely gone aesthetic. While browsing a favorite corner of the internet not long ago, I ran into this work by designer/generative artist **Manolo Gamboa Naon**



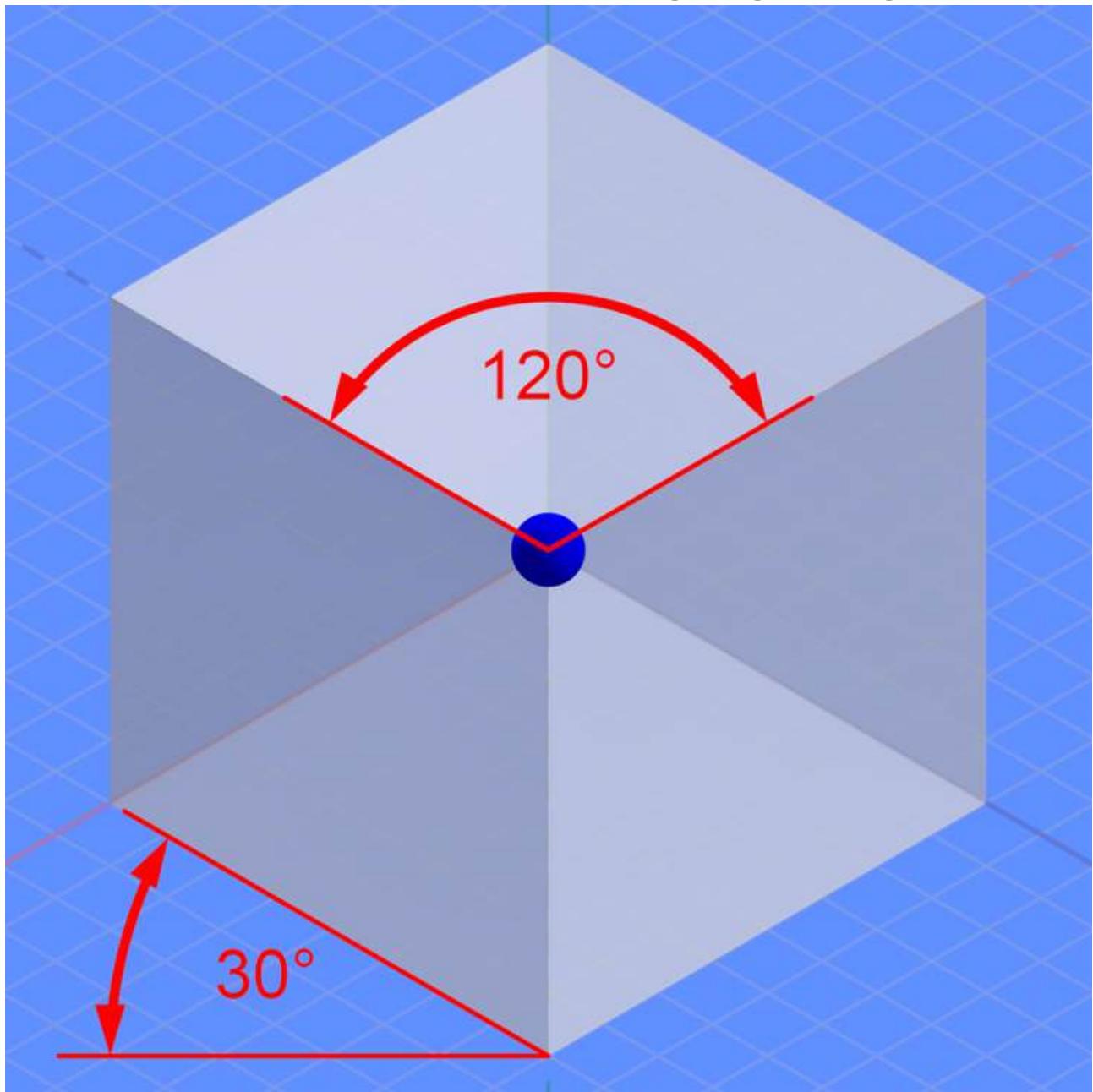
I'm a sucker for this sort of thing. My first thought was: how can this concept be reproduced with a little data behind it? Not enough data to produce the burden of "analysis." Not enough to straighten it out and kill its beauty. Just enough to seed the design. After sprinkling some MLB stats over a pile of smoke and mirrors, I had this:



**Job done...** but before I could move on to The Next Thing, CJ asked me to share the technique with y'all. I hope to do that below, with a much simpler set of data.

## ISOMETRICS

The purpose of an isometric perspective is to display multiple sides of an object as equally as possible. This necessarily means viewing an object centered on a corner, with each visible face getting 120 degrees of view:



Credit: By SharkD – Own work. Download source code., CC BY-SA 4.0,  
<https://commons.wikimedia.org/w/index.php?curid=70888147>

This creates an impression of 3D space that is both clinically precise and slightly unreal. There is no horizon and the isometric space scrolls forever in every direction. Features are neither larger nor smaller with respect to “distance” from the viewer.

Video games have long taken advantage of isometric/similar perspectives, to gain the advantages of having a third dimension but with a locked camera. Think *Project Zomboid*, *Bastion*, or the *Civilization* and *Sims* franchises. You can even buy isometric graph paper. How cool is that.

What follows may not produce a geometrically pure isometric perspective, but since it's close and "isometric adjacent" is a clumsy phrase at best, let's just call it isometric from here on out.

## THE PATH

We are going to follow these steps:

1. Polygon Squares
2. Isometric "Top" Faces
3. All Faces

## POLYGON SQUARES

We will start with a simple data set and create five squares, using polygons. If you are new to polygons in Tableau then you're in luck, as CJ has already produced three detailed tutorials on how to build and manipulate them. You can view the series here, [pt1](#), [pt2](#), [pt3](#).

To super summarize the rules of polygons for our purposes:

Points: every point needs a set of coordinates: an x (conventionally on columns) and a y (conventionally on rows).

Lines: a line simply connects two points. That means one line needs two data rows.

Polygons: a polygon is a series of lines resulting in a closed figure. Lines holding hands.

We are going to logically join two tables (dataset is in the repo!) : our content which is a list of television characters, and a scaffold to build the squares. Each character has been given an origin coordinate:

Abc subjects character	# subjects size	# subjects x_origin	# subjects y_origin
Earn	2	2	2
Darius	2	4	2
Van	2	0	4
Alfred	4	8	2
Lottie	1	4	0

Followed with the scaffolding,

# quad'scaffold point_id
1
2
3
4
5

and the relationship between them.

The screenshot shows the Tableau Data Source interface. At the top, there are two boxes: 'characters' on the left and 'quad\_scaffold' on the right, connected by an orange arrow pointing from 'characters' to 'quad\_scaffold'. Below this, a dropdown menu shows 'characters — quad\_sca...'. A tooltip or help text below the dropdown says 'How do relationships differ from joins? Learn more'. Below the dropdown, there are three fields: 'characters' with value "'this is fine'", 'Operator' with value '=', and 'quad\_scaffold' with value "'this is fine'". There is also a button labeled '(+) Add more fields'.

The scaffold is intentionally devoid of any meaningful content. Its only job is to generate five rows per character. Each row will represent one point, and—wait, what's that? Squares only have four corners but our scaffold has five rows? Yes. In order for Tableau to properly close a polygon, its final point must have identical coordinates to its first point. That means an “extra” row in the data. I like to start my polygons at lower left, and work clockwise.

For visual variety, we will use our subjects’ field to determine polygon sizes. Here are the calcs for our x and y coordinates:

```

x_poly
// Builds the x coordinate for each point
// for each character.
// Starts at lower left and proceeds clockwise.
CASE
WHEN 1 THEN
WHEN 2 THEN
WHEN 3 THEN +
WHEN 4 THEN +
WHEN 5 THEN
END
    
```

```

        y_poly
// Builds the y coordinate for each point
// for each character.
// Starts at lower left and proceeds clockwise.

```

```

CASE
WHEN 1 THEN
WHEN 2 THEN +
WHEN 3 THEN +
WHEN 4 THEN
WHEN 5 THEN
END

```

Because our calcs return numeric values, Tableau assumes they are measures. To prevent Tableau from trying to aggregate them, we will make them continuous dimensions instead: right-click each calc to “Convert to Dimension” and then right-click again to “Convert to Continuous”.

Here is a crosstab proving the output for Darius:

character	size	point_id	x_poly	y_poly
Darius	2	1	4	2
		2	4	4
		3	6	4
		4	6	2
		5	4	2

Now we plot:

Put on Detail and on Color.

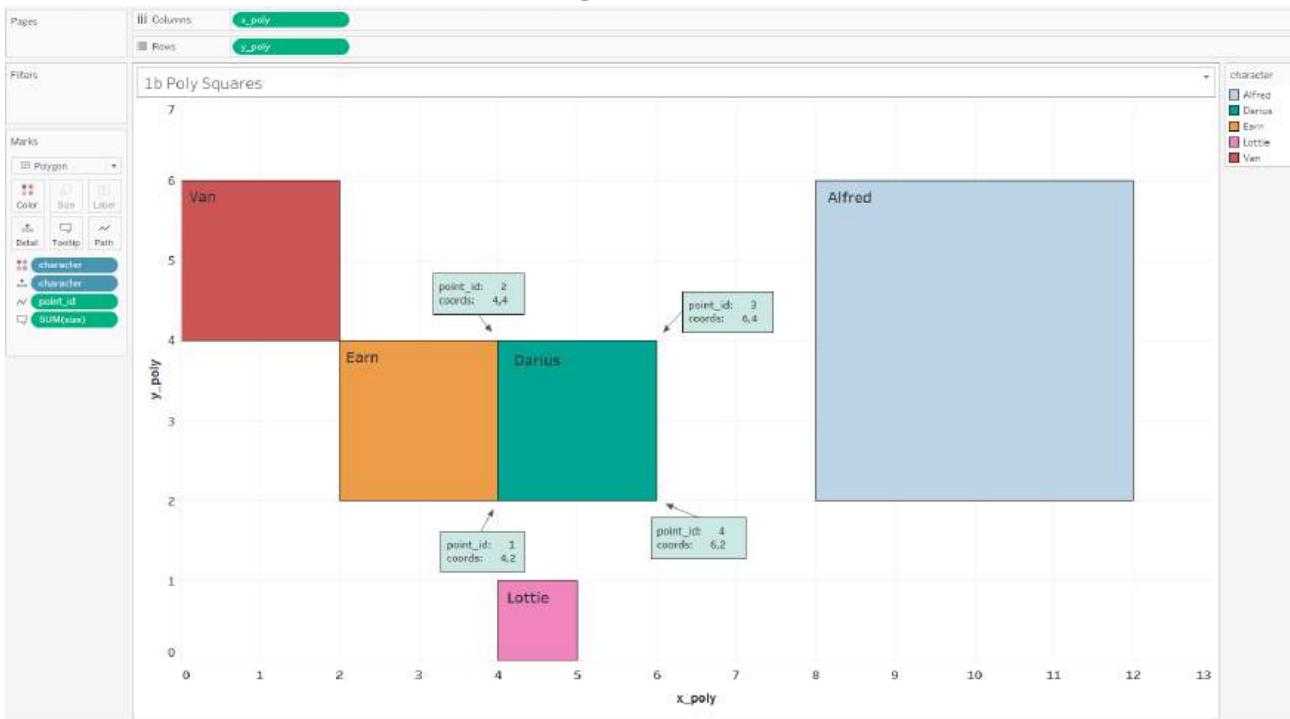
Put on Columns.

Put on Rows.

Change the mark type from “Automatic” to “Polygon”. Once Tableau knows we are trying to build a poly it unlocks the Path card, which is its way of saying, “in what order shall I connect these dots?”

Put on Path. Because that's the correct order.

All in all, we have told Tableau: “Draw one polygon per character. Start at the coordinates for point 1, and keep going through point 5.” Which gives us



## ISOMETRIC “TOP” FACES

Are there mathematically rigorous ways to convert a square face into an isometric face? Almost certainly! Being neither mathematical nor rigorous, I found a Dumb Guy Math solution. It’s basically the ole Spin’n’Squish:

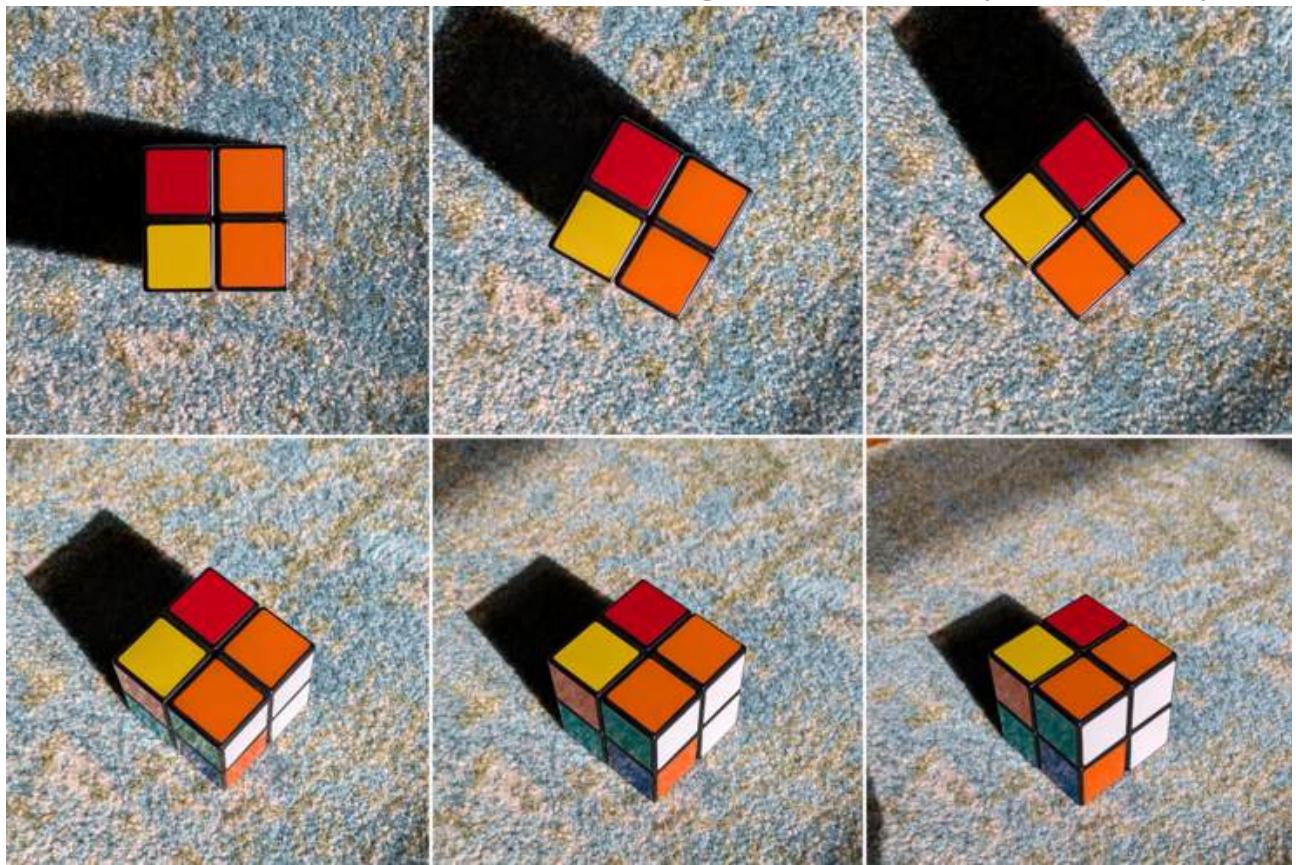
Spin the object 45 degrees

Squish the vertical

We are going to call these calcs and as a mental trick for what happens going forward. I mean, you have seen thousands of bar charts and line charts and scatterplots by now. Have you ever thought about your viewing perspective of those charts? Are you looking “down” at the bars, or from the “side”? It doesn’t matter of course; this is not a useful concept in flatland.

But to help our brains keep track of the transforms to come, let’s pretend we have been looking down at the “tops”, and now we are going to change our point of view. As demonstrated with, um, this Rubik’s cube

on the rug in my office. First we rotate our perspective (the “spin”), then we lower the camera until achieving isometric-ism (the “squish”)



(Mega bonus points to anyone who knows which movie inspired me to buy this thing, *which no I cannot solve*)

```
x_iso_top  
// clockwise:  
// +
```

```
// counter-clockwise:
```

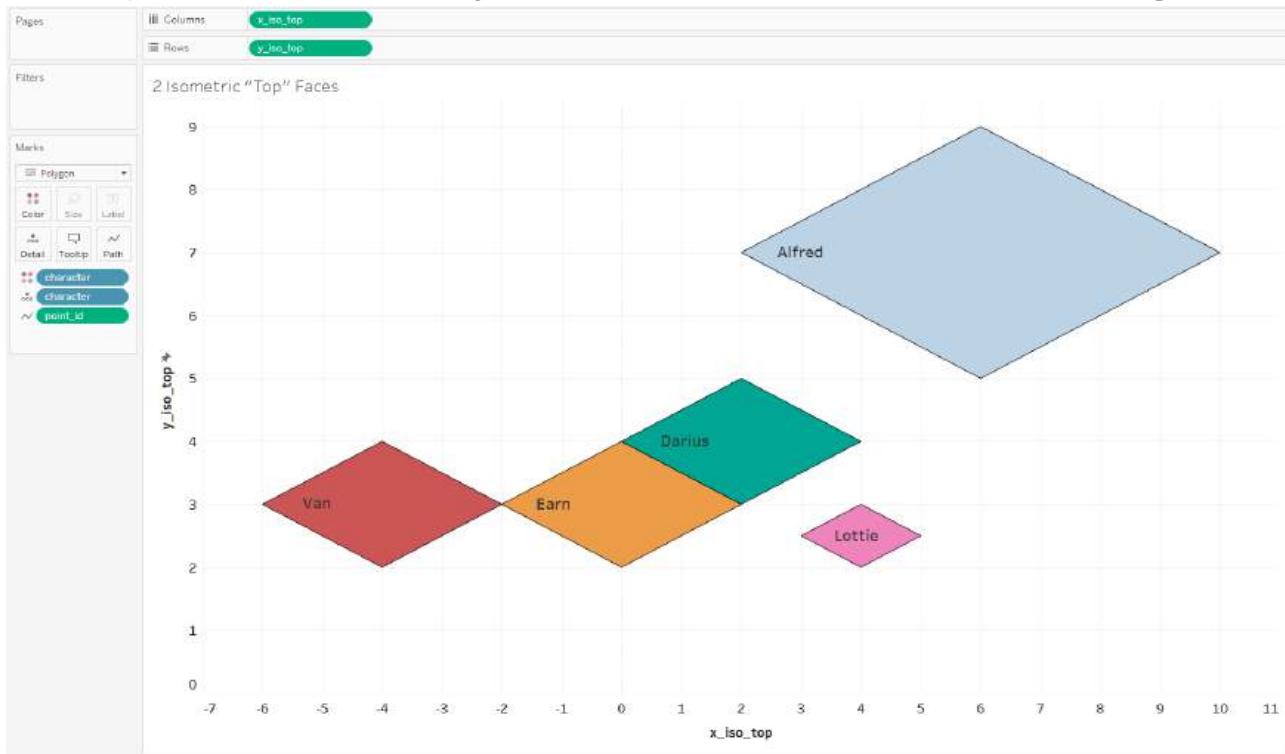
```
-
```

```
y_iso_top  
// clockwise:  
//( - ) / 2
```

```
// counter-clockwise:  
( + ) / 2
```

Convert these calculations to continuous dimensions as before.

We replace the old x and y coordinates with our new calcs, and get this



Your brain may register these objects as two-dimensional diamonds. In flatland your brain is correct. But remember the Rubik's cube!

### ALL FACES

To complete the cubes, we have one more data structure trick, and it's the same trick as before: another scaffold. The first scaffold generated five rows per character, so that we could make a four-sided polygon. This new scaffold will multiply that effect. Now we will have three faces per character, and five points per face. The data source & relationship below



[How do relationships differ from joins? Learn more](#)



Drag the new scaffold onto the canvas and create a relationship with the first scaffold, again with a full join.

Here's the cool part. There is no need to calculate the isometric coordinates for the front faces. The fronts are just “hanging” off of the top. Every necessary x coordinate is already present in the top face. And the y coordinates are either a) touching the top face, or b) a known distance directly below that top face.

To simplify the downstream work, we write a few calcs to identify the x and y coordinates of selected points on the top face.

```
x_iso_top_pt_1
{ FIXED :
MIN(IF = 1 THEN END) }
```

```
x_iso_top_pt_2
{ FIXED :
MIN(IF = 2 THEN END) }
```

```
x_iso_top_pt_4  
{ FIXED :  
MIN(IF = 4 THEN END) }
```

```
y_iso_top_pt_1  
{ FIXED :  
MIN(IF = 1 THEN END) }
```

```
y_iso_top_pt_2  
{ FIXED :  
MIN(IF = 2 THEN END) }
```

```
y_iso_top_pt_4  
{ FIXED :  
MIN(IF = 4 THEN END) }
```

Now that we have isolated the coordinates for the top face, we can apply them to build out our fronts.

As a poor SQL writer, I love me some nested conditionals. Daunting at first glance, these calcs simply check which face and which point is to be assigned a coordinate, and then generate that value. Did you think these points would be determined by a single elegant calculation or something? Who am I, Bora Beran?

```
x_iso_all
CASE
WHEN 'top' THEN
WHEN 'left' THEN
CASE
WHEN 1 THEN
WHEN 2 THEN
WHEN 3 THEN
WHEN 4 THEN
WHEN 5 THEN
END
WHEN 'right' THEN
CASE
WHEN 1 THEN
WHEN 2 THEN
WHEN 3 THEN
WHEN 4 THEN
WHEN 5 THEN
END
END
```

```
y_iso_all
CASE
WHEN 'top' THEN
WHEN 'left' THEN
CASE
WHEN 1 THEN -
WHEN 2 THEN
WHEN 3 THEN
WHEN 4 THEN -
WHEN 5 THEN -
END
WHEN 'right' THEN
CASE
WHEN 1 THEN -
WHEN 2 THEN
WHEN 3 THEN
WHEN 4 THEN -
WHEN 5 THEN -
END
END
```

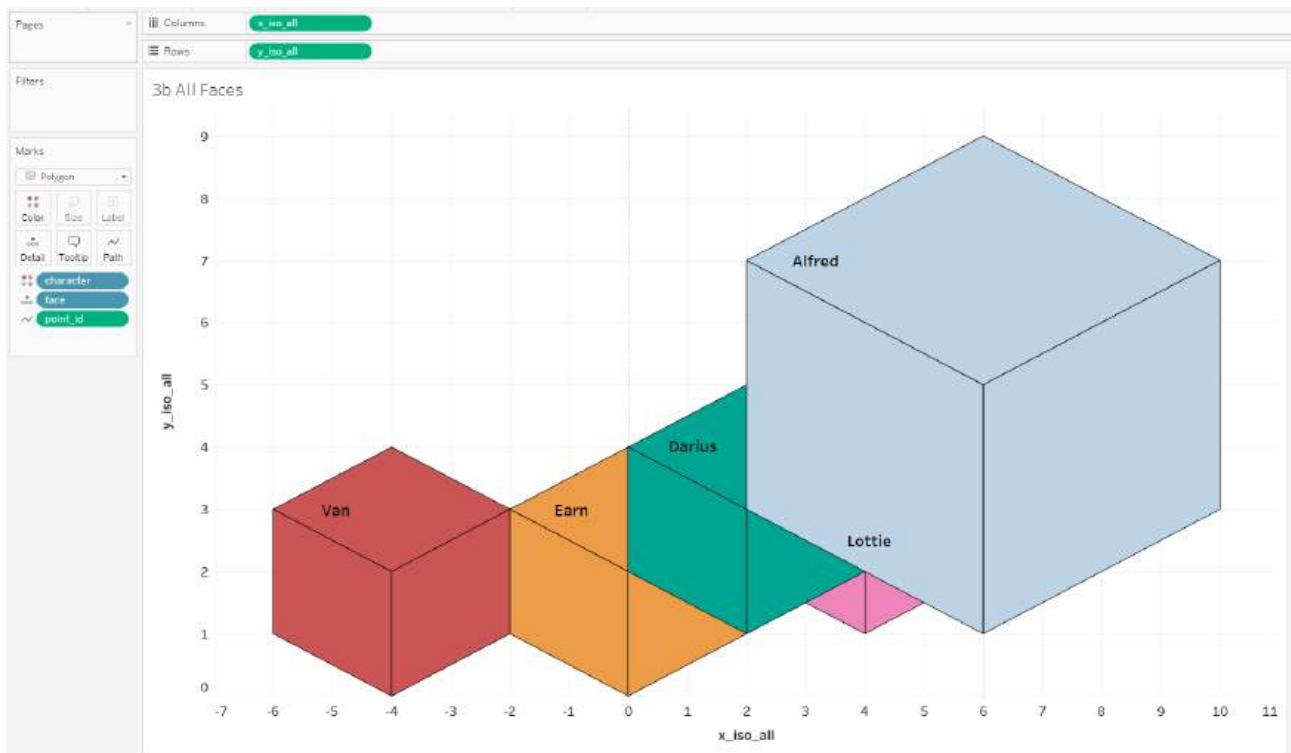
Here is Darius again, now with all three faces and five points.

character	face	point_id	x_iso_all	y_iso_all
Darius	left	1	0	2
		2	0	4
		3	2	3
		4	2	1
		5	0	2
	right	1	2	1
		2	2	3
		3	4	4
		4	4	2
		5	2	1
	top	1	2	3
		2	0	4
		3	2	5
		4	4	4
		5	2	3

Once again, convert them to continuous dimensions. Once again, drag these new x and y calcs to replace the last set: to columns and to rows.

Add to Detail, because now each poly is being drawn at the - grain.

And we have... a mess!



This piece of lazy Escher art is brought to you by default sorting, in this case ascending alpha by . Alpha makes “Alfred” in front of any other

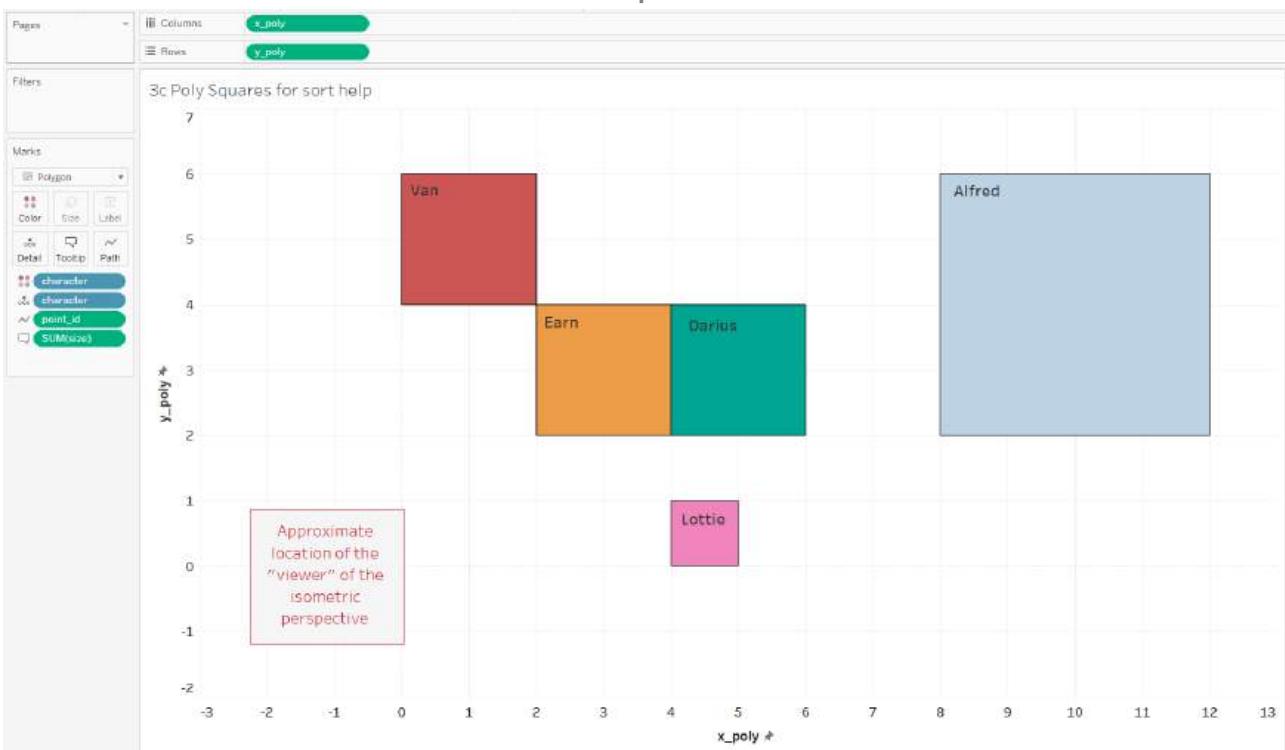
character. Argh.

Sorting is a challenge. While working on my original piece (which had almost 600 MLB players and therefore almost 1,800 polygon faces), I spent literally half of my time on task trying to keep Adam Frazier in front

of Ezequiel Tovar. It was maddening and in the end I cheated by manually building a sort field in my data source. The more objects and

the denser their clustering, the more likely you will run into such an obstacle, and require a creative solution. Dumb Guy Math will not save us.

Fortunately our simple example offers a simple solution. Let's look at the original squares again, with a new annotation suggesting our isometric viewpoint.



From that viewpoint, Darius should be “in front” of Alfred and both Lottie and Earn should be “in front” of Darius. An obvious sorting rule presents itself: “smaller Y values should appear to be in front of larger Y values. As tiebreaker, smaller X values should appear to be in front of larger X values.” We can encode these values into a new field

```
sorter  
STR()  
+ ';'  
+ STR()
```

(BONUS QUESTION: what would happen if any of our characters had any origin coordinate less than 0 or greater than 9? How would you solve that?)

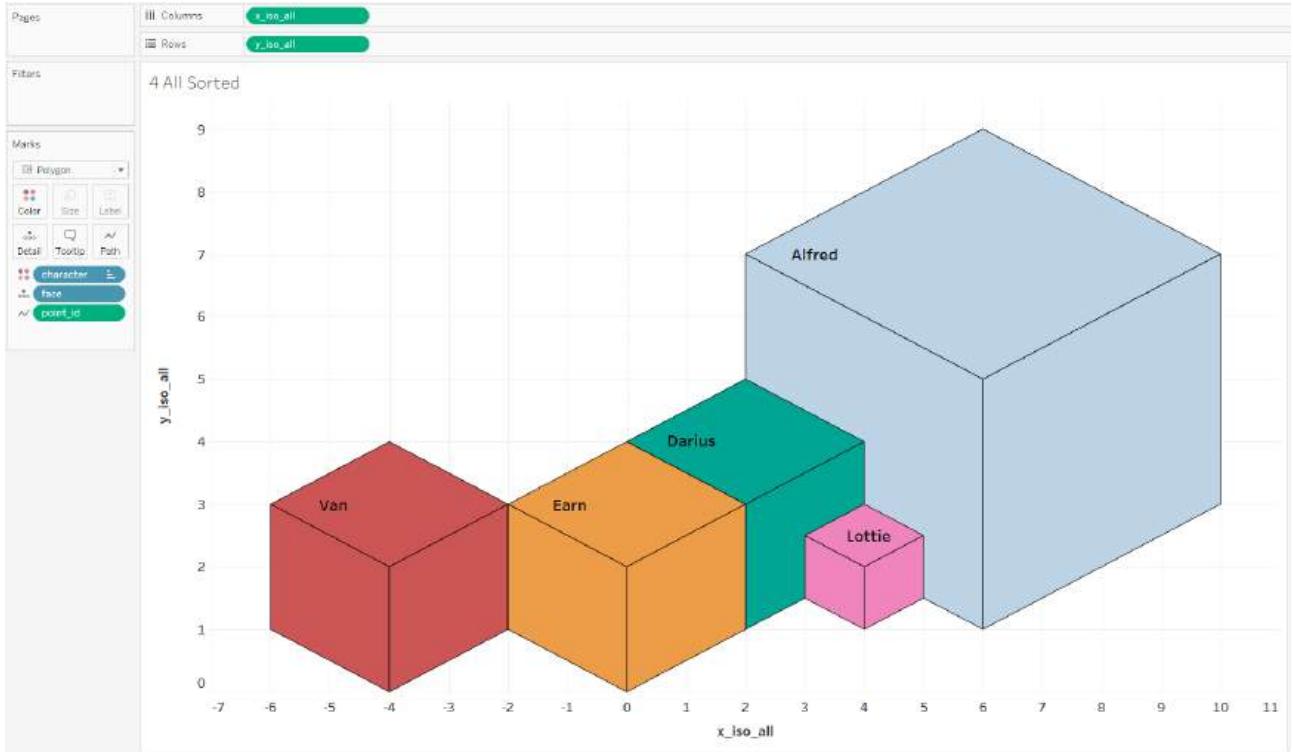
Sort by the field aggregated by “Minimum,” and lazy Escher is banished.

The screenshot shows the Tableau interface with the Sort dialog box open. The Sort dialog box has the following settings:

- Sort By:** Field
- Sort Order:** Ascending (selected)
- Field Name:** sorter
- Aggregation:** Minimum

On the left, the Marks shelf shows three selected fields: character (highlighted in blue), face (highlighted in blue), and point\_id (highlighted in green).

Which leaves us with this,



That's better. The view is now complete from a basic technical standpoint. This is a nice example of how an isometric view is nicely clinical yet slightly unreal, and how this unrealism may create perceptual problems for the viewer. My brain reads this as though all of the cubes are sitting on top of the same plane. It also sees Darius abutting Alfred, with Lottie snuggled up in the corner there. Neither perception is correct, but there is not enough visual information to convince my brain otherwise.

My original viz employed some further cosmetic tweaks that could be applied here. Each face is a distinct polygon and can be colored independently of the others. We can provide some spacing between cubes by proportionally reducing face size. Instead of drawing the “top” faces on the same plane, we can put the “bottom” edges on the same plane. And fifty other things I haven’t thought of, but you might.

### SHOULD WE EVEN?

I am not the Viz Police. The most casual of glances at my Tableau Public profile will attest to that. We could easily fill a bathtub with Edward Tufte quotes to discourage/condemn applications of our little isometric exercise. Ducks. Lie factors. All legit, in their own context. But there is

one Tufte quote that has been sticking in my head this whole time. From the introduction to *Envisioning Information*:

*"The world is complex, dynamic, multidimensional; the paper is static, flat. How are we to represent the rich visual world of experience and measurement on mere flatland?"*

I like that.

Counterpoint: You know who needs a 3D bar chart? Nobody does.

CJ Round-up:

Love it! A really articulate walkthrough of the needed scaffolding to create isometric shapes and also the sorting required to make it come to life with a 3D plane. If you would like a copy of the template Matt has walked through you can find it on my profile, along with the git repository for the data.

## ISOMETRIC CHARTS IN TABLEAU WITH MATT ANTOLINE

03/02/2025



[Return to category](#)

If you end up using the logic please be sure to credit Matt's original viz on MLB! I've already got some ideas cooking.

LOGGING OFF,

CJ

## UNDERSTAT SOCCER ETL PROCESS WITH JORDAN PICKLES

Hi all,

Recently I got the chance to chat with Jordan who has been posting some pretty awesome projects on his Git Repository and Tableau Public.

One that caught my eye was in the football community.

Today he was kind enough to share some of the data processing that goes into retrieving, cleaning and storing understat data in a database. A great project to replicate for those in the football community or looking to expand some of their core analytics skills.

I'm excited to see how Jordans work expands, he has some serious talent and such a clear method to the way he works and presents that information back in such an articulate manner, just check out the blog below.

Jordan, over to you.

### ETL Process – From API to Database

Hi, I'm Jordan Pickles, a data analyst currently working in the aviation and travel industry. My academic background is in Exercise and Health Scientific Research, where I contributed to clinical trials and Health Tech R&D projects. Along the way, I discovered how much I loved working with data, which led me to make the jump from research into a data role.

To develop my skills, I completed the AI Core Data Science Accelerator program, which gave me a solid technical foundation. For the past two years, I've been putting those skills to work as a data analyst, turning data into insights and solving real-world problems.

Over the past few months, I've spent some of my spare time working with football data, aiming to improve my data skills as well as working with some different and intriguing data. In November 2024, I submitted my first visualization to SportsVizSunday, which extracted data from the

Understat API package and visualised the *Top Goalscorers in European Soccer Leagues in Europe* while also developing techniques for dynamic zone visibility in tableau ([see blog post](#)).

## Top Premier League Goalscorers in the 2024/25 Season

Click on a bar to expand / collapse player details

1 Erling Haaland

**11 Goals | Manchester City**

2 Chris Wood

**8 Goals | Nottingham Forest**

**Chris Wood**

Type of Shot

Goal Scored From

Goals vs xG:

RightFoot 5

DirectFreekick

Goals 8

Head 3

FromCorner 1

xG 5.4

OtherBody...

OpenPlay 3

Variance +2.6

LeftFoot

Penalty 1

SetPiece 1



3 Bryan Mbeumo

**8 Goals | Brentford**

4 Mohamed Salah

**7 Goals | Liverpool**

5 Cole Palmer

**7 Goals | Chelsea**

6 Danny Welbeck

**6 Goals | Brighton**

7 Nicolas Jackson

**6 Goals | Chelsea**

8 Ollie Watkins

**5 Goals | Aston Villa**

9 Luis Díaz

**5 Goals | Liverpool**

10 Liam Delap

**5 Goals | Ipswich**

Data: Understat.com | Creator: Jordan Pickles

Building on that project I wanted to dive deeper into handling data in Python by creating a fully automated ETL (Extract, Transform, Load) pipeline resulting in a database consisting of a match data table and shot data table which can be queried for projects going forward.

Detailed below is the process of extracting data from the Understat API package, cleaning and transforming the data in Python and loading the data into a PostgreSQL database. All the code for this project can be found in the [GitHub Repository](#).

### Data Collection, Wrangling, and Cleaning

Understat.com provides data for football matches across the top 5 european leagues and the UnderstatAPI package provides a user-friendly endpoint to the API, making it an excellent starting point for football data analysis. The package provides options to retrieve information at various levels of aggregation (e.g. match, shot, teams etc). For further details on the possibilities of the Understat API package, check out the [documentation](#).

## Match Data Extraction

The first step when working with the understat API package is to initiate the client (see Figure 1 below) which provides functions to interact with the API. Once the client has been initiated the next step is to collect the unique IDs for each match across the top five European leagues. The code below iterates through a list of league names, retrieving data for all matches in the league for the 2024 season by using the `get_match_data` function.

```
1 import understatapi
2
3 league_list = ['EPL', 'La_Liga', 'Bundesliga', 'Serie_A', 'Ligue_1']
4 season = '2024'
5
6 class UnderstatDataScraper:
7     def __init__(self):
8         self.client = understatapi.UnderstatClient()
9
10    def match_id_retrieval(self, league_list, season, max_date):
11        match_id_list = []
12        match_data_list = []
13
14        for league in league_list:
15            league_matches =
self.client.league(league=league).get_match_data(season=season)
```

*Figure 1. This code Initialises the Understat API client and collects a list of Match ID's for each of the leagues available in the 2024 season.*

The function returns match data for each league in a semi-structured JSON format (see Figure 2 below) which is not suitable for analysis / uploading to a database, therefore the data needs to be extracted.

```
[{"id": "28048", "isResult": True, "h": {"id": "282", "title": "Le Havre", "short_title": "LHV"}, "a": {"id": "161", "title": "Paris Saint Germain", "short_title": "PSG"}, "goals": {"h": "1", "a": "4"}, "xG": {"h": "0.197023", "a": "2.9098"}, "datetime": "2024-08-16 18:45:00", "forecast": {"w": "0.0045", "d": "0.0384", "l": "0.9571"}}, {"id": "28049", "isResult": True, "h": {"id": "241", "title": "Brest", "short_title": "BRE"}, "a": {"id": "164", "title": "Marseille", "short_title": "MAR"}, "goals": {"h": "1", "a": "5"}, "xG": {"h": "1.51475", "a": "2.96496"}, "datetime": "2024-08-17 15:00:00", "forecast": {"w": "0.085", "d": "0.1596", "l": "0.7554"}}]
```

*Figure 2. Shows the semi-structured data output for just two of the games retrieved from the code block in Figure 1.*

With the data in a semi-structured format, the next task required is to extract the data. The code in Figure 3 below iterates through each match present in the JSON output for the league currently being iterated through. The code initially checks that the game has already taken place (excluding future matches) and ensures that the match date is later than the latest date already in the database (when this isn't the first time the code has been run), so only new games are returned. Once the match has been determined as required to be added to the database, the code extracts the necessary data into a dictionary of key-value pairs (e.g. Key: datetime: Value: 01/01/2025 15:00:00) with the data types declared for each variable. Each dictionary, representing one match, is then added to a list, until all matches from all leagues are present in the list. Once all the matches are present in the list it can be converted into a pandas DataFrame providing a structured table of columns and rows which is suitable for analysis / uploading to a database.

```

16 # This is continuation of the match_id_retrieval function in Figure 1
17     for match in league_matches:
18         try:
19             game_date = datetime.strptime(match['datetime'], '%Y-%m-%d %H:%M:%S')
20             today = datetime.today()
21             # Finds the games that are not already in the database to optimise the code
22             if game_date < today and game_date > max_date and match['goals']['h'] != None:
23                 match_id_list.append({'id': match['id'], 'league': league})
24
25 # Parses through the data returned and appends the data to a list which is converted to a df later in the code
26     match_data_list.append({
27         'match_id': int(match['id']),
28         'datetime': datetime.strptime(match['datetime'], '%Y-%m-%d %H:%M:%S'),
29         'league': str(league),
30         'home_team_id': int(match['h']['id']),
31         'home_team_name': str(match['h']['title']),
32         'home_team_name_short': str(match['h']['short_title']),
33         'away_team_id': int(match['a']['id']),
34         'away_team_name': str(match['a']['title']),
35         'away_team_name_short': str(match['a']['short_title']),
36         'home_team_goals': int(match['goals']['h']),
37         'away_team_goals': int(match['goals']['a']),
38         'home_team_xg': float(match['xG']['h']),
39         'away_team_xg': float(match['xG']['a']),
40         'win_forecast': float(match['forecast']['w']),
41         'draw_forecast': float(match['forecast']['d']),
42         'loss_forecast': float(match['forecast']['l']),
43         'season': int(season)
44     }
45 )
46     except KeyError as e:
47         print(f"KeyError: {e} in match: {match}")
48
49 df_all_match_data = pd.DataFrame(match_data_list)
50

```

*Figure 3. This code iterates through each Match present in the league\_matches variable declared in Figure 1, extracting the information from each match returned in that league before adding the dictionary to a list which is later converted into a data frame. Note: This code is a continuation of the function in Figure 1.*

### Shot Data

A list of unique match ID's collected from the code in Figure 3 can now be used to call the `get_shot_data` function, as shown in Figure 4 below.

This function also returns data in a semi-structured format, with one dictionary for the home team data and one for the away team data. The code below extracts both dictionaries separately before concatenating them on top of each other to create one data frame for each match. All data frames are then appended to a list which can then be converted into one big data frame at the end of the code block. Similar to how the match data frame was created.

```

1 def match_shots(self, match_id_list, season):
2     match_dataframes = []
3     for id in match_id_list:
4         try:
5             match_data = self.client.match(id['id']).get_shot_data()
6             df_home = pd.DataFrame(match_data['h'])
7             df_away = pd.DataFrame(match_data['a'])
8
9             # Concatenates the home and away data for this match
10            df_match = pd.concat([df_home, df_away], ignore_index=True)
11            df_match['league'] = id['league']
12            df_match['season'] = season
13            match_dataframes.append(df_match)
14        except Exception as e:
15            print(f"Error: {e} in match: {id}")
16            df_shot_data = pd.concat(match_dataframes,
17 ignore_index=True)
18    return df_shot_data

```

*Figure 4. This code iterates through the match ID list collected in the Figure 1 and Figure 3 and calls the get\_shot\_data function of the Understat API client to return the shot data for each match ID which is extracted and concatenated into a one data frame.*

### Cleaning the Shot Data

Thanks to accessing the API, the data is well-structured and needs little cleaning. However, during exploratory data analysis and when converting the text data (the output of the Understat API function) to the appropriate data types, a few areas required some additional cleaning.

The pandas and numpy packages in Python offer a range of functions that simplify the data cleaning process. In this project, the data cleaning involved tasks such as replacing parts of strings in columns, renaming and dropping columns, adding new columns based on specific logic, converting columns to the appropriate data types, dropping rows without coordinate data present and rescaling the coordinate data for the intended use case. The X and Y coordinates from Undertstat are between 0 & 1 but in-order to work on the images I visualise the data on (Statsbomb pitch of the mplsoccer package), an axis of 120, 80 is required, therefore the X and Y coordinates are rescaled.

```

1 def clean_shot_data(self, df_shot_data):
2     #Removes any apostrophe's from the column
3     for col in ['X', 'Y', 'xG']:
4         df_shot_data[col] = df_shot_data[col].astype(str).str.replace("'", "", regex=False)
5
6     # Renames the columns to more readable names
7     df_shot_data = df_shot_data.rename(columns={
8         'id':'shot_id', 'shotType':'shot_type', 'lastAction':'last_action', 'xG':'xg', 'X':'x', 'Y':'y',})
9
10    # Finds the team of the player taking the shot and then drops the columns not required
11    df_shot_data['player_team'] = np.where(df_shot_data['h_a'] == 'h', df_shot_data['h_team'],
12                                         df_shot_data['a_team'])
13    df_shot_data['team_against'] = np.where(df_shot_data['h_a'] == 'h', df_shot_data['a_team'],
14                                         df_shot_data['h_team'])
15    df_shot_data = df_shot_data.drop(columns=['h_team', 'a_team', 'h_goals', 'a_goals'])
16
17    df_shot_data = df_shot_data.dropna(subset=['X', 'Y'])
18
19    # Converts the data types of the columns to the correct data types
20    df_shot_data = df_shot_data.astype({
21        'shot_id': 'int', 'minute': 'int', 'result': 'str', 'x': 'float', 'y': 'float', 'xg': 'float',
22        'playerstr', 'h_a': 'str', 'player_id': 'int', 'situation': 'str', 'season': 'int', 'shot_type': 'str',
23        'last_action': 'str', 'player_team': 'str', 'player_assisted': 'str', 'date': 'datetime64[ns]',
24        'leaguestr', 'team_against': 'str',
25        'match_id': 'int'})
26
27    #Re-scales coordinates suitable to use on pitch maps
28    df_shot_data['x'] = df_shot_data['x']*120
29    df_shot_data['y'] = df_shot_data['y']*80
30
31    return df_shot_data

```

*Figure 5. This code Cleans the shot df\_shot\_data frame that has been created in the previous code blocks.*

### Database Creation and Management

The previous steps have now provided two clean and structured data frames suitable for use in the primary objective of this project, to build a database storing the collected data that can be queried not just for this project but for future projects as well. For the database server, I chose PostgreSQL ([download here](#)), which is a free database server that can be downloaded to both macOS and Windows. To interact with PostgreSQL from Python, there are two main packages: Psycopg2, a native PostgreSQL package in python, and SQLAlchemy, a Python package that supports interaction with various databases, including PostgreSQL.

Psycopg2 allows for a direct connection to the Database API (DBAPI) and the execution of SQL commands as strings. On the other hand,

SQLAlchemy provides Python-specific functions and objects that simplify interacting with the database (e.g. functions that are converted into SQL dialect by the package). In this project, I used a combination of both packages, leveraging the strengths of each to build and manage

the database efficiently. For further reading into the differences between SQLAlchemy and PostgreSQL, check out this [article](#).

### Step 1 – Installing the Packages

To begin, the two required packages need installing, using a package manager of choice. In this project, I used pip as the package manager, see below the pip commands required in the terminal:

- Psycopg2: ‘`pip install psycopg2`’
- SQLAlchemy: ‘`pip install SQLAlchemy`’

### Step 2 – Connecting to the Database

#### *Connecting with Psycopg2*

Establishing a connection to the DBAPI from inside of your code editor is the key step, this allows you to interact with the database whilst still having all of the previous steps / variables in the code available to use, call and upload.

Psycopg2 connects to the PostgreSQL DBAPI using the `psycopg2.connect(<<Connection String>>)` function (documentation for further detail). The connection contains multiple keyword argument details relevant to the local database which can be seen below (see figures 6 & 7). In this case, I’ve stored these database credentials in a YAML file, which can be easily opened and read with Python.

```
1 # Set up the YAML File with details
2 LOCAL_HOST: LOCALhost
3 LOCAL_PASSWORD: <<Your Password>>
4 LOCAL_USER: postgres
5 LOCAL_DATABASE_TYPE: postgresql
6 LOCAL_DATABASE: <<Your Database Name>>
7 LOCAL_PORT: 5432
8 LOCAL_DB_API: psycopg2
```

*Figure 6. YAML file containing the database credentials required for the Psycopg2 connection.*

The `psycopg2_connect` function in my code (see Figure 7 below) handles the scenario where the database has not yet been created. If the database named in the YAML file doesn't already exist, a connection to the server is created instead of the database. The `create_new_db` function is then called to create the database within the PostgreSQL server. Once a connection has been established, a cursor is created, which allows Python code to execute PostgreSQL commands in a database session, just as is required when managing or querying the database. This `create_new_db` function creates a new database with the name stored in the YAML file using the cursor before closing the connection to the server, then the `pscyopg2_conenction` function can connect directly to the database.

```
1 import psycopg2
2 from psycopg2 import sql
3 import os
4 import yaml
5
6 def read_db_creds(self):
7     with open("Credentials.YAML", "r") as f:
8         db_creds = yaml.safe_load(f)
9     return db_creds
10
11 def psycopg2_connect(self, db_creds):
12     # Connects to the database if created
13     try:
14         psycopg2_connection = psycopg2.connect(
15             dbname = db_creds['LOCAL_DATABASE'],
16             user = db_creds['LOCAL_USER'],
17             password = db_creds['LOCAL_PASSWORD'],
18             host = db_creds['LOCAL_HOST'],
19             port = db_creds['LOCAL_PORT']
20         )
21     except Exception as e:
22         #Connects to PSQL Server
23         psycopg2_connection = psycopg2.connect(
24             user = db_creds['LOCAL_USER'],
25             password = db_creds['LOCAL_PASSWORD'],
26             host = db_creds['LOCAL_HOST'],
27             port = db_creds['LOCAL_PORT']
28         )
29         self.create_new_db(psycopg2_connection,
30 db_creds['LOCAL_DATABASE'])
31     # Connects to target DB
32     psycopg2_connection = psycopg2.connect(
33         dbname = db_creds['LOCAL_DATABASE'],
34         user = db_creds['LOCAL_USER'],
35         password = db_creds['LOCAL_PASSWORD'],
36         host = db_creds['LOCAL_HOST'],
37         port = db_creds['LOCAL_PORT']
38     )
39     print("Connected to the database successfully!")
40     return psycopg2_connection
41
42 def create_new_db(self, psycopg2_connection, new_db_name):
43     try:
44         # Set to autocommit ensuring the action is implemented
45         psycopg2_connection.autocommit = True
46         with psycopg2_connection.cursor() as cursor:
47             cursor.execute(sql.SQL("CREATE DATABASE {}").
48                             format(sql.Identifier(new_db_name)))
49         psycopg2_connection.close()
50         print(f"Database '{new_db_name}' created successfully!")
51
52     except Exception as e:
53         print(f"Error creating database '{new_db_name}': {e}")
54
```

**Figure 7.** This code reads the credentials required for the database from the YAML file in figure 6 and uses the details to connect to the database with Psycopg2. If the database does not already exist then the `create_new_db` is called to create the database.

### Connecting with SQLAlchemy

Connecting with SQLAlchemy requires an additional step than the Psycopg2 connection, this is creating an engine before establishing the connection to the database. The engine allows dialect with the DBAPI, in essence to translate the SQLAlchemy python objects / functions into SQL commands that are given to the DBAPI and in turn, actioned on the database.

To create the engine, a string containing the necessary connection details is passed through the `create_engine(<<Connection String>>)` function (see documentation). Once the engine is created, the `engine.connect()` function can be called to establish the connection to the DBAPI.

```
1 def init_db_engine(self, db_creds):
2     try:
3         # Create the engine for the target database
4         engine = create_engine(f"{{db_creds['LOCAL_DATABASE_TYPE']}+"
5         {{db_creds['LOCAL_DB_API']}://{{db_creds['LOCAL_USER']}:{db_creds['LOCAL_PASSWORD']}@{{db_creds['LOCAL_HOST']}:"}
6         {{db_creds['LOCAL_PORT']}/{{db_creds['LOCAL_DATABASE']}}}")
7         connection = engine.connect()
8         print("Connected to the database successfully!")
9         return engine, connection
10    except Exception as e:
11        print(f"Error initializing database engine: {e}")
```

**Figure 8.** Code to create the SQLAlchemy engine and connect to the database.

### Step 3 – Creating the Tables

This is where the power of SQLAlchemy really shines. With SQLAlchemy, you can create database tables, add constraints, map Python data types to those of your chosen database server and establish relationships between tables.

In the GitHub repository for this project, you'll find a Python file called `Models.py`. Inside this file, there are three classes (see Figure 9 below):

- `class Base(DeclarativeBase)`
- `class Match(Base)`
- `class Shot(Base)`

The `Base(DeclarativeBase)` class is a base class that allows for mapping Python data types to their corresponding PostgreSQL data types and is inherited by other classes such as when defining table schemas, in this case for the *Match* and *Shot* tables. Each class specifies details like column names, data types, and relationships to other tables. A primary key of a table provides a unique ID (e.g. an ID for each shot) and a foreign key provides a typically non-unique ID that can be used to join tables on (e.g. match id that each shot occurred in for the shot table). Typically to ensure that joins are successful, the data should be joined on a primary key from one table and a foreign key of another table. As can be seen in the table schemas for this project, the primary key of the `dim_match` table (Match ID, a unique value) has a relationship with the foreign key of the `dim_shots` table (Match ID, non-unique value).

```

1 from sqlalchemy import Column, Integer, String, Float, Date, DateTime, ForeignKey
2 from sqlalchemy.orm import declarative_base, relationship, Mapped, mapped_column, Relationship,
   DeclarativeBase
3 from sqlalchemy.dialects.postgresql import TEXT, VARCHAR, INTEGER, FLOAT, DATE, TIMESTAMP
4 from typing import List
5 from typing import Optional
6
7 class Base(DeclarativeBase):
8     type_annotation_map={
9         int: INTEGER,
10        str: VARCHAR,
11       float: FLOAT,
12      Date: DATE,
13     DateTime: TIMESTAMP
14    }
15
16 class Match(Base):
17     __tablename__ = 'dim_match'
18
19     match_id: Mapped[int] = mapped_column(Integer, primary_key=True)
20     datetime: Mapped[DateTime]
21     league: Mapped[str] = mapped_column(String(10))
22     home_team_id: Mapped[int]
23     home_team_name: Mapped[str] = mapped_column(String(30))
24     home_team_name_short: Mapped[str] = mapped_column(String(3))
25     away_team_id: Mapped[int]
26     away_team_name: Mapped[str] = mapped_column(String(30))
27     away_team_name_short: Mapped[str] = mapped_column(String(3))
28     home_team_goals: Mapped[int]
29     away_team_goals: Mapped[int]
30     home_team_xg: Mapped[float]
31     away_team_xg: Mapped[float]
32     win_forecast: Mapped[float]
33     draw_forecast: Mapped[float]
34     loss_forecast: Mapped[float]
35     season: Mapped[int]
36     shots: Mapped[List["Shot"]] = Relationship(back_populates="match")
37
38 class Shot(Base):
39     __tablename__ = 'dim_shot'
40
41     shot_id: Mapped[int] = mapped_column(Integer, primary_key=True)
42     minute: Mapped[int]
43     result: Mapped[str] = mapped_column(String(30))
44     x: Mapped[float]
45     y: Mapped[float]
46     xg: Mapped[float]
47     player: Mapped[str] = mapped_column(String(50))
48     h_a: Mapped[str] = mapped_column(String(1))
49     player_id: Mapped[int]
50     situation: Mapped[str] = mapped_column(String(20))
51     season: Mapped[int]
52     shot_type: Mapped[str] = mapped_column(String(20))
53     last_action: Mapped[str] = mapped_column(String(20))
54     player_team: Mapped[str] = mapped_column(String(30))
55     player_assisted: Mapped[str] = mapped_column(String(50))
56     date: Mapped[Date]
57     league: Mapped[str] = mapped_column(String(10))
58     team_against: Mapped[str] = mapped_column(String(30))
59     match_id: Mapped[int] = mapped_column(ForeignKey('dim_match.match_id'))
60     season: Mapped[int]
61     match: Mapped["Match"] = Relationship(back_populates="shots")

```

*Figure 9. This is the models.py file declaring the base class which is inherited when defining the table schemas for the dim\_match and dim\_shot tables defined in a separate class. The constraints and relationships of the tables are also declared.*

Once the table schemas are defined, the `Base.metadata.create_all(bind=engine)` function is called to create the

tables in the database, if they do not already exist.

#### Step 4 – Loading the Data

Now that the data has been extracted and transformed, the final stage is loading the data into the database. This step can be performed using either Psycopg2 or SQLAlchemy. However, the process differs slightly between the two:

With Psycopg2, data is loaded by iterating / looping through each row in the DataFrame and inserting them row by row using the *INSERT SQL* command.

On the other hand, SQLAlchemy offers a more efficient method with its native *to\_sql()* function. This function allows you to insert the entire DataFrame at once, with several parameters to control the action (e.g. append, update, or replace the existing table). In this project, I opted to *append* the data, as my pipeline is designed to only return data for games not already in the database, thus optimising the time taken to run the process on a weekly basis etc.

For performance reasons, I chose to insert the data in batches of 1,000 rows at a time (see Figure 10 below). This helped avoid issues when uploading large data frames (first time running the code) and improved the overall efficiency of the process.

```
1 def append_db_table(self, data_frame, table_name, engine):
2     batch_size=1000
3     try:
4         data_frame.to_sql(table_name, engine, if_exists='append', index=False,
5                            chunksize=batch_size)
6         print(f"Data appended to table {table_name} successfully!")
7     except Exception as e:
8         print(f"Error appending data to table {table_name}: {e}")
```

*Figure 10. This function takes in the table name and the relevant data frame ready to be loaded into the table. The data is appended to the table if it already exists to ensure only the latest data collected in the code previously is added to the data table.*

#### The Database

The aim of this project was to create a fully automated ETL pipeline resulting in a database consisting of a match data table and shot data

table which can be queried for projects going forward. The steps outlined in this article have resulted in a PostgreSQL database consisting of two clean, 2-dimensional tables (see figure 11 below for an example of the dim\_shot table).

	shot_id [PK] integer	minute integer	result character varying (30)	x double precision	y double precision	xg double precision	player character varying (50)
1	584630	19	MissedShots	116.51999816894532	53.840002441406256	0.050078392028808594	Bruno Fernandes
2	584631	26	MissedShots	89.63999633789062	38.320001220703126	0.020807035267353058	Casemiro
3	584632	28	SavedShot	109.56000366210938	27.679998779296877	0.25769174098968506	Bruno Fernandes
4	584633	33	SavedShot	109.56000366210938	37.039999389648436	0.4701629877090454	Bruno Fernandes

*Figure 11. Example of the dim\_shot table created in the database.*  
The benefit of having a database is that you can query and wrangle the data into any format that you need in an easy and time efficient way using SQL. Queries can be written directly and returned in PgAdmin (Front-end application for PostgreSQL) or through Psycopg2 using python in your code editor (a simple example can be seen in figure 12 below).

```
1 cursor = self.connection.cursor()
2 query = sql.SQL("SELECT MAX(datetime) FROM dim_match") # Gets the max date from the database
3 tabler.execute(query)
4 max_date = cursor.fetchone()[0]
5 cursor.close()
```

*Figure 12. This code shows how a database can be queried with a Psycopg2 cursor connection. This code returns the first row of the SQL string to find the max date in the dim\_match table.*

This article has shown how to leverage the understAPI client to extract the data, transform the data and then load the data into a database that can be used for many analyses in the future.

A big thank you to CJ for asking me to feature on the site which provides some brilliant content for analysts and has certainly been a help to me.

Please feel free to reach out to me on [linkedin](#) if you have any questions or wish to connect.

CJ Roundup:

I think a lot of community members probaby are use to retrieving data through various packages but to then start to think about the transformation and storage of the data, to then visualise is what has really impressed me with Jordans work. Make sure to take a look at his

repo to take full advantage of learning the steps he went through.

Jordan, thank you for making time.

LOGGING OFF,

CJ

## BUILDING HABITS WITH MY KINDLE

Hi all,

I hope you have settled into the new year. The turn of the year often comes with a reflective period. I tweeted at the start of the year a few of my goals which included;

- 30 mins x 330 days exercise
- 20 mins x 330 days kindle
- Run another ultra marathon
- 1 fitness holiday
- Grow site and sports youtube content
- More time with gf & dog
- Alcohol free

Some of these things are easy to measure, and some are also harder to track than others.

Reading has always been a tough one for me, I just don't enjoy it compared to other things I could be doing with the time, but I also know what great benefits it can bring to my vocabulary, my focus, and my relaxation. So this year, I wanted to make sure I read my kindle as much as possible, (other wise whats the point in paying for kindle unlimited right?)

if you'd like a copy of the code written you can find it in the repository at the top of the page.

If you're new here, each blog is also replicated in resource on my page.

The screenshot shows a GitHub repository interface. The repository name is 'Website-Tutorials / 2025 Blogs'. The left sidebar lists files categorized by year: 2021, 2022, 2023, 2024, and 2025. Under 2025, there are several blog posts: '2025-01-02 Tableau Wrapped ...', '2025-01-22 Building Habits WI...', 'DS\_Store', and 'Readme.md'. The main content area shows a commit history with 11 entries, the most recent being 'Update Readme.md' at 17 minutes ago. A note at the bottom of the repository page states: 'Within this repository, you will find details of every blog I posted in 2025 on cjmayer.com'.

Today we will see how far short of my goal I am, for 20 mins of reading my kindle for 330 days this year.  
Accessing Amazon Data

The screenshot shows the 'Request your data' page on the Amazon UK website. The URL is [amazon.co.uk/hz/privacy-central/data-requests/preview.html](https://amazon.co.uk/hz/privacy-central/data-requests/preview.html). The page features a search bar and navigation links for All, Amazon Basics, Today's Deals, Morrisons, Buy Again, Gift Cards & Top Up, Gift Ideas, Browsing History, cjm's Amazon, Customer Service, Fashion, Prime, Prime Video, Books, and Best Sellers. Below the navigation is a link to 'Amazon.co.uk - Today's Deals - Retail - Outlet - Subscribe & Save - Vouchers - Amazon Prime - Prime Video - Prime Student - Mobile Apps - Amazon Pickup Locations'. The main content area is titled 'Request your data' and includes a note: 'Please select the data that you want. Don't forget you can access the majority of your data instantly, as well as update your personal information, from Your Account.' A dropdown menu labeled 'Select Data Category' is shown, with 'Kindle' selected. A 'Submit Request' button is at the bottom, along with a note: 'You can access the majority of your data instantly, as well as update your personal information, from Your Account. If you have a more specific request or for further support, please Contact Us.' There is also a link to 'Learn more about how we collect and use data to provide and improve our services.'

To begin, export your Amazon purchase history. The specific URL varies by country (e.g., different for Amazon US and Amazon UK). If you're in another country, simply use your local Amazon domain. Once there, select the "Kindle" category and click "Submit Request." Depending on your Kindle usage, it might take several days to receive the download. For me it was overnight. It's worth noting you can actually download a whole bunch of data beyond just kindle from Amazon!

You will end up with a zip file to download. Unzip that folder. There will be about 74 folders within that. Let's take a look at just two for now, book relations and reading session.

First lets read in a few files.

## The ones I think are of value are

- Content Updates (Mainly because it had the title of the books in, there are other files that also do, but this had the whole list for me) I don't end up using the book name, but it may be useful for your own analysis.

```
▶ ✓ 4 minutes ago (<1s) 7
%python
# Define a dictionary where keys are the original column names and values are the new names
columns_to_keep = {
    'Product Name': 'Product Name',
    'ASIN': 'ASIN'
}

# Keep only the specified columns, rename them, and get distinct rows
df_br = df_br[list(columns_to_keep.keys())].rename(columns=columns_to_keep).drop_duplicates()

# Display the result
display(HTML(df_br.head().to_html()))



|   | Product Name                                                                                                                                                                                   | ASIN       |
|---|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------|
| 0 | Ikigai & Kaizen: The Japanese Strategy to Achieve Personal Happiness and Professional Success (How to set goals, stop procrastinating, be more productive, build good habits, focus, & thrive) | B0CN3NFSXK |
| 2 | Habit Stacking: 97 Small Life Changes That Take Five Minutes or Less                                                                                                                           | B00JQHB67O |
| 3 | Dopamine Detox : A Short Guide to Remove Distractions and Get Your Brain to Do Hard Things                                                                                                     | B098MHBF23 |


```

- ReadingSession – This is the really valuable file as it has the start and end times to all the kindle reads. Start & End times as well as page flips are captured here.

```
▶ ✓ 12:15 PM (<1s) 12
%python
# Assuming your DataFrame is named 'df' and the column with milliseconds is named 'duration_ms'
df_rs['duration_minutes'] = df_rs['total_reading_millis'] / (1000 * 60)
df_rs['start_timestamp'] = pd.to_datetime(df_rs['start_timestamp'], errors='coerce')
df_rs = df_rs[df_rs['start_timestamp'].dt.year == 2025]

▶ ✓ 12:15 PM (<1s) 13
%python
# Define a dictionary where keys are the original column names and values are the new names
columns_to_keep = {
    'start_timestamp': 'start_timestamp',
    'end_timestamp': 'end_timestamp',
    'ASIN': 'ASIN',
    'duration_minutes': 'duration_minutes',
    'number_of_page_flips': 'number_of_page_flips'
}

# Keep only the specified columns and rename them
df_rs = df_rs[list(columns_to_keep.keys())].rename(columns=columns_to_keep)

# Display the result
display(HTML(df_rs.head().to_html()))



|    | start_timestamp           | end_timestamp        | ASIN       | duration_minutes | number_of_page_flips |
|----|---------------------------|----------------------|------------|------------------|----------------------|
| 4  | 2025-01-04 09:35:52+00:00 | 2025-01-04T09:45:24Z | B074LZG7KS | 9.533333         | 23.0                 |
| 6  | 2025-01-09 22:34:03+00:00 | 2025-01-09T22:46:27Z | B074LZG7KS | 11.401667        | 33.0                 |
| 10 | 2025-01-02 23:18:52+00:00 | 2025-01-02T23:19:19Z | B06XC4125R | 0.453333         | 11.0                 |
| 12 | 2025-01-16 22:10:49+00:00 | 2025-01-16T22:15:56Z | B0CN3NFSXK | 5.126667         | 15.0                 |
| 23 | 2025-01-14 21:23:28+00:00 | 2025-01-14T21:36:57Z | B00JQHB67O | 13.481667        | 39.0                 |


```

As part of the code i'll do a little clean up too, adding flexibility for renaming columns as well as limiting it just to this years books. We then join those two files together to get a start and end time, the number of pages flipped and the title of the book.

You'll see recently the book I read on stoicism.

▶ ✓ 12:15 PM (<1s)

```
%python
# Display the result
display(HTML(df_titled.head().to_html()))
```

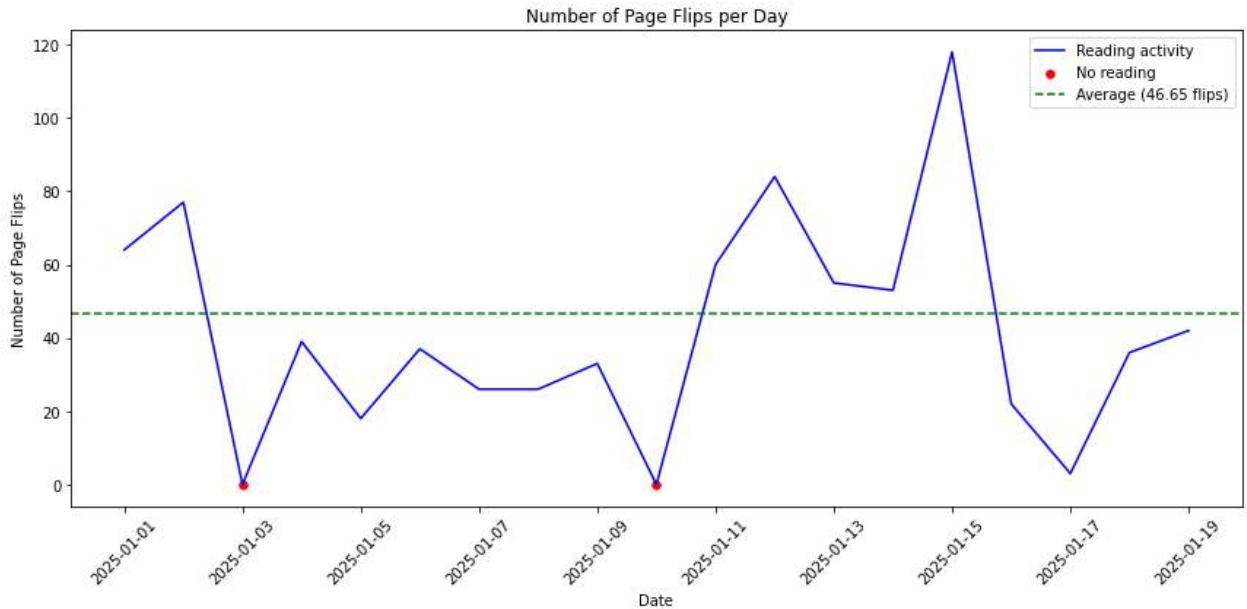
	start_timestamp	end_timestamp	ASIN	duration_minutes	number_of_page_flips	Product Name
35	2025-01-01 09:20:18+00:00	2025-01-01T09:20:22Z	B086RGHJ2K	0.055000	1.0	Stoicism: How to Use Stoic Philosophy to Find Inner Peace and Happiness
29	2025-01-01 09:20:25+00:00	2025-01-01T09:29:40Z	B086RGHJ2K	9.258333	26.0	Stoicism: How to Use Stoic Philosophy to Find Inner Peace and Happiness
9	2025-01-01 09:35:13+00:00	2025-01-01T09:40:27Z	B086RGHJ2K	5.235000	12.0	Stoicism: How to Use Stoic Philosophy to Find Inner Peace and Happiness
41	2025-01-01 20:53:19+00:00	2025-01-01T21:01:32Z	B086RGHJ2K	8.206667	16.0	Stoicism: How to Use Stoic Philosophy to Find Inner Peace and Happiness
31	2025-01-01 21:10:06+00:00	2025-01-01T21:14:20Z	B086RGHJ2K	4.243333	9.0	Stoicism: How to Use Stoic Philosophy to Find Inner Peace and Happiness

With that light prep we can obviously now start to look at what it means for my goal. There are a few questions i have in mind.

- Am I reading every day?
- Am i even reading 20 minutes a night? I don't think I am. I feel like i am rushing my read.
- I read alot of non fiction.... I tend to just skim read the books, skipping chapters or pages where I already know the concept. How does this impact the page turns a day.



Quickly we see the number of pages flips a day is reasonable.... but it doesn't highlight the days I missed nor gives me a sense of how many pages a day this way. Also what happened on the 16th? Did i just want to mark a book complete?

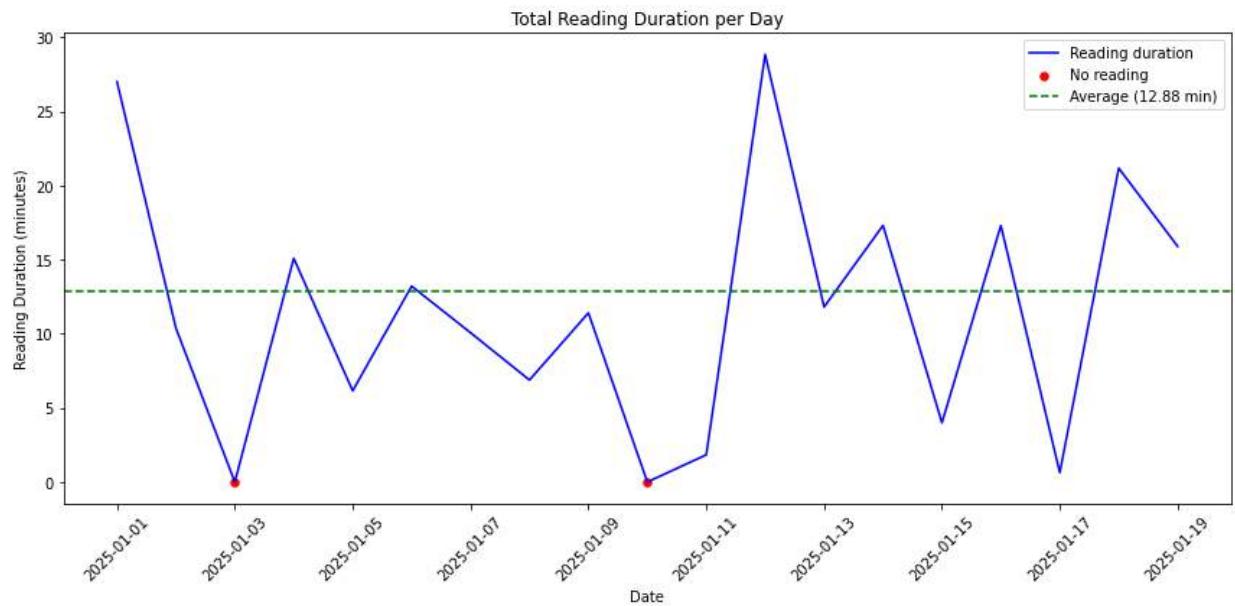


We can add a little more context through filling in the date gaps making it continuous and bringing light to the days i didn't read. Interestingly enough the 3rd of Jan is my birthday... so i'll let myself off. On the 10th Jan I started reading at 3 minutes past midnight. Shame!

The chart shows i flip over 40 pages a day. But in truth this will be lower in terms of actual read for two reasons.

1. The scan reading technique
2. Flicking back and forth to re read pages.

So let's close and see how truthful my actual reading duration target is. We use the same continuous date range, and roll up the read time each night, excluding days with no read from the average.



Okay.... thats not as bad as i thought. 13 minutes on average.

It does highlight a few things for me though.

- I'm doing okay at loading the kindle, I just need to really relax into reading it. It might be worth removing other distractions when I go to read so that I can solely focus on reading.
- Skim reading is okay, but perhaps I need to go beyond what I normally read around self-help, business, fitness and health to spend more time digesting the information on each page.

Cheers to spending more time building better habits. Some reading is better than no reading.

LOGGING OFF,

CJ

#### TABLEAU WRAPPED WITH HARI JEYAN

Hi all,

A happy new year! I hope 2025 brings you many great things.

The turning of the year is always a great time to reflect on the year previous and plan for the year ahead.

You may have seen at the end of last year a “Tableau Wrapped” doing the circuits in the data community built by Hari Jeyan. I was so impressed, and it’s had an incredible amount of praise from those looking to get some juicy numbers about their portfolio. Today, I had the chance to sit down with Hari to talk through how he went about creating Tableau Wrapped, the spin off of the popular spotify wrapped from years back.

## CJ's Tableau Public Journey

Tableau Public



Hello,

**CJ Mayes**

You're publishing Workbooks since,

**5 years, 5 months and 6 days**

It's a journey of perseverance



You've published

**204**

Tableau Workbooks



You're an inspiration to

**2,136**

people who follow you



You follow

**800**

people who inspire you



You've accumulated

**165,071**

views on your workbooks



Your workbooks have

**3,960**

favorites from users

**#TableauWrapped**

Created by: Hari Jeyan

Generated on: 02/01/2025

CHECK OUT THE VIDEO FROM HARI



## HOW DID I GET INTO DATA?

Thank you, CJ, for encouraging me to share the story and lessons from a simple Tableau Wrapped web app I built recently. You can check it out here: [Tableau Wrapped](#).

I'm **Hari Jeyan**, from the beautiful country of Sri Lanka, and I've been using Tableau since early 2022. Although I hold a Bachelor's degree in Information Technology, which involved extensive coding, I realized during my internship in 2020 as a front-end developer that my true interests lay elsewhere.

Three weeks into my software engineering internship, I transitioned to a Business Analyst role, as I felt more drawn to the business side of the company rather than coding. This shift sparked my curiosity about finding a career that combined my technical skills with business decision-making.

Eventually, I discovered my passion for working at the intersection of business and data, where I could use technology to drive data-informed decisions. Towards the end of my degree, I took courses on Coursera and started experimenting with basic Tableau Public projects. However, apart from these initial projects and my professional work, I wasn't actively engaging with the Tableau community or consistently building data analytics projects to showcase my skills.

## THE PROJECT IDEA

Although I only started using Spotify in 2024, I was amused by their Spotify Wrapped feature since they started. It was incredible to see how this simple yet personalized concept became a powerful marketing catalyst, inspiring similar initiatives across various tools and platforms. It was the evening of December 20th, a Friday, and my social media feed was flooded with Spotify Wrapped, LinkedIn Wrapped, and other year-end summaries. This end-less posts of Wrapped features sparked a series of questions in my mind:

- Why haven't data professionals shared anything Wrapped-related for their portfolios?
- Does Tableau have a Wrapped feature?
- Why doesn't Tableau offer something like this for its vast community?
- What would it take to build a Tableau Wrapped?
- Can I do it?
- Will my perfectionist mindset drag this project out forever?
- Should I build this in public?

- What if I fail?

With just ten days left until the New Year, I was in between the excitement of starting this project and the fear of not finishing it in time.

Ultimately, I decided to move forward for one reason: I believed this could make a difference for the Tableau datafam community.

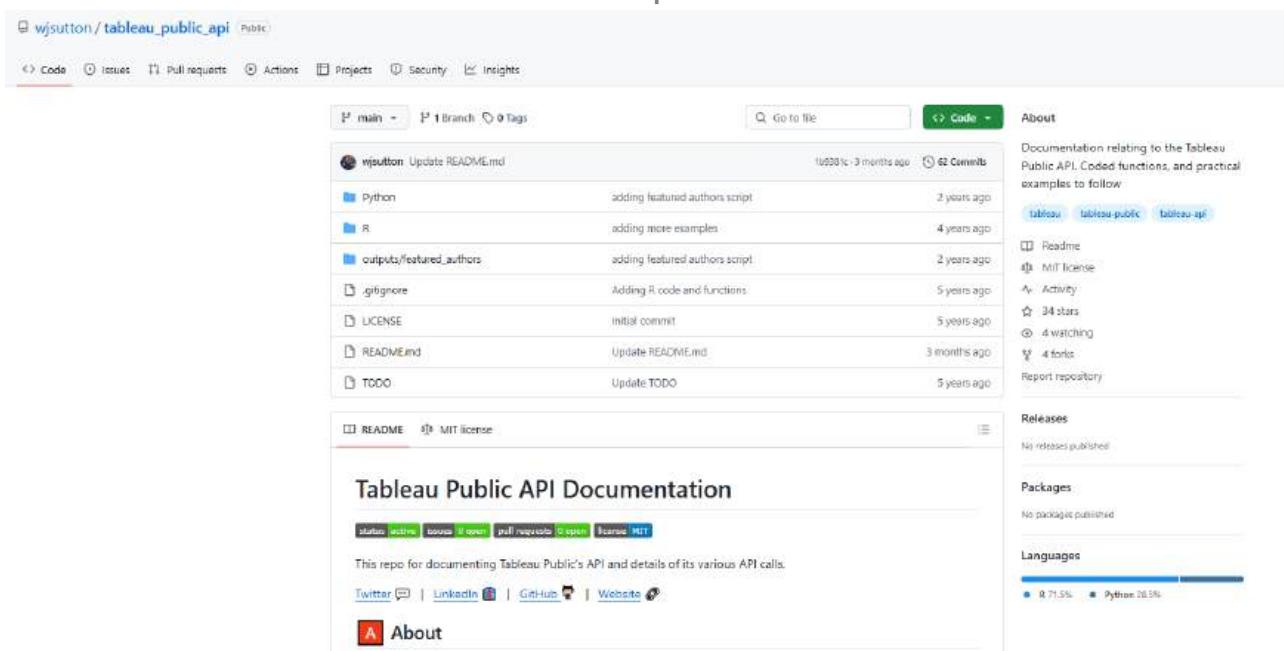
All I needed to do was create something that people would love to use.

## RESOURCES

It had been more than four years since I last wrote code to build something, aside from data wrangling. I initially saw this as a major roadblock. However, I realized that the Tableau Public API, which is well-documented by Will Sutton

([https://github.com/wjsutton/tableau\\_public\\_api](https://github.com/wjsutton/tableau_public_api)), was readily available.

This meant I could focus on making API requests and using the responses without having to start from scratch like building a web scraper.



A few weeks before, I watched a YouTube video titled “How to Build Side Projects with AI?” by Harish Uthayakumaran and his team. They explained how they built a GMAT practice web app in just a few days using AI tools. It was a lightbulb moment for me. Their experience showed that building products today is never like before. AI tools make the process faster and more accessible. Inspired by this, I wanted to dive into building something myself, but at that point, I wasn’t sure what to create.

With this new perspective, I felt less concerned about the coding aspect of the project. Instead, I shifted my focus to figuring out the UI, deciding on meaningful metrics, and addressing other key elements needed to bring my idea to life.

## TABLEAU PUBLIC API

The success of this entire project is dependent on one crucial component: the Tableau Public API. This API provides access to data about user profiles and workbooks that are publicly shared on the Tableau Public platform, making it the backbone of the project.

API call format: <https://public.tableau.com/profile/api/> + Tableau Public Username

Name	X	Headers	Preview	Response	Initiator	Timing	Cookies
↳ ip.json?referrer=https%3A%2F...		▼ General					
✖ author							
□ bg9s?x-amz-cf-id=NFMlucVUi...							
↳ hari.jeyan							
▣ collect?v=2&tid=G-8YLN0SNX...							
↳ workbook?excludeHiddenWor...							
↳ hari.jeyan							
↳ workbooks?profileName=hari.j...		▼ Response Headers					
↳ workbooks?profileName=hari.j...							
▣ collect?v=2&tid=G-8YLN0SNX...							
↳ announcement_banner							

## Profile API Response:

X	Headers	Preview	Response	Initiator	Timing	Cookies
▼ {profileName: "hari.jeyan", name: "Hari Jeyan", title: "Data Analytics Engineer", organization: "",...}						
avatarUrl: "https://public.tableau.com/avatar/e54a2b08-c36c-42a3-8e4e-fc332af3bdaf.jpeg"						
bio: ""						
createdAt: 1651363200000						
featuredVizRepoUrl: "StoryofEmiratesAirlines"						
freelance: true						
hideNewWorkbooks: false						
name: "Hari Jeyan"						
organization: ""						
profileName: "hari.jeyan"						
profileView: "DEFAULT"						
searchable: true						
title: "Data Analytics Engineer"						
totalNumberOfFollowers: 9						
totalNumberOfFollowing: 74						
visibleWorkbookCount: 24						
▼ websites: [{id: 1975065, url: "https://www.linkedin.com/in/harijeyan/", title: "linkedin.com"},...]						
▶ 0: {id: 1975065, url: "https://www.linkedin.com/in/harijeyan/", title: "linkedin.com"}						
▶ 1: {id: 1975066, url: "https://tableau-wrapped.vercel.app/", title: "website1"}						
▶ 2: {id: 1975064, url: "https://twitter.com/Harijeyan_08", title: "x.com"}						

The following are the API objects returned in the response. Here we can find some of the useful data that can be used in the wrapped.

1. **avatarUrl**: The URL of the user's profile picture/avatar.
2. **bio**: The biography or description of the user.
3. **createdAt**: The account creation timestamp in milliseconds since the Unix epoch.
4. **featuredVizRepoUrl**: The URL or identifier for the user's featured visualization repository.
5. **freelance**: A boolean indicating whether the user is available for freelance work.
6. **hideNewWorkbooks**: A boolean indicating whether new workbooks are hidden from the user's profile.
7. **name**: The full name of the user.
8. **organization**: The organization the user is associated with.
9. **profileName**: The user's profile name or username.
10. **profileView**: The view setting for the profile (e.g., "DEFAULT").
11. **searchable**: A boolean indicating whether the profile is searchable.
12. **title**: The professional title or designation of the user.
13. **totalNumberOfFollowers**: The total number of followers the user has.
14. **totalNumberOfFollowing**: The total number of other profiles the user is following.
15. **visibleWorkbookCount**: The number of workbooks visible on the user's profile.

### Tableau Public API URL for Workbooks

X	Headers	Payload	Preview	Response	Initiator	Timing	Cookies
▼ General							
Request URL:				https://public.tableau.com/public/apis/workbooks?profileName=harijeyan&start=0&count=12&visibility=NON_HIDDEN			
Request Method:				GET			
Status Code:				200 OK			
Remote Address:				13.225.4.126:443			
Referrer Policy:				origin-when-cross-origin			

API call format for first 50 workbooks:

[https://public.tableau.com/public/apis/workbooks?profileName= + Tableau Public Username + &start=0&count=50&visibility=NON\\_HIDDEN](https://public.tableau.com/public/apis/workbooks?profileName= + Tableau Public Username + &start=0&count=50&visibility=NON_HIDDEN)

The key is to utilize the start and count parameters effectively, incrementing them by 50 with each request, as the API retrieves workbook details in batches of 50!

```

X Headers Payload Preview Response Initiator Timing Cookies
▼ {current: 0, next: 12,...}
  ▼ contents: [{workbookRepoUrl: "StoryofEmiratesAirlines", title: "Story of Emirates Airlines",...},...]
    ▶ 0: {workbookRepoUrl: "StoryofEmiratesAirlines", title: "Story of Emirates Airlines",...}
      authorDisplayName: "Hari Jeyan"
      authorProfileName: "hari.jeyan"
      defaultViewName: "Emirates"
      defaultViewRepoUrl: "StoryofEmiratesAirlines/sheets/Emirates"
      numberOffavorites: 1
      title: "Story of Emirates Airlines"
      viewCount: 150
      workbookRepoUrl: "StoryofEmiratesAirlines"
    ▶ 1: {workbookRepoUrl: "UserEngagementAnalysis-Infinity", title: "User Engagement Analysis - Infinity",...}
    ▶ 2: {workbookRepoUrl: "2020-Test1-0129", title: "Organization Efficiency", defaultViewName: "Story 1",...}
    ▶ 3: {workbookRepoUrl: "LineChartsLegends", title: "Line Charts & Legends",...}
    ▶ 4: {workbookRepoUrl: "DOAC_YT_Analytics", title: "DOAC_YT_Analytics", defaultViewName: "DOAC KPI",...}
    ▶ 5: {workbookRepoUrl: "London_Bike_Rides", title: "London Bike Rides", defaultViewName: "Dashboard 1",...}
    ▶ 6: {workbookRepoUrl: "SuperstoreSuperDashboard", title: "Superstore Super Dashboard",...}
    ▶ 7: {workbookRepoUrl: "Dubai20-minMobility", title: "Dubai 20-min Mobility",...}
    ▶ 8: {workbookRepoUrl: "Dendrogram_or_TreeChart",...}
    ▶ 9: {workbookRepoUrl: "Covid-19TrackerbyCountry_16556338760800", title: "Covid-19 Tracker by Country",...}
    ▶ 10: {workbookRepoUrl: "CovidData_16546148928850", title: "Covid Statistic Tracker",...}
    ▶ 11: {workbookRepoUrl: "AnalyzingDonaldTrumpsTweets_16545845693900",...}
  current: 0
  next: 12

```

Here also we can see some of the useful data that are specific to individual dashboards.

1. **authorDisplayName:** The display name of the workbook's author.
2. **authorProfileName:** The profile name (or username) of the workbook's author.
3. **defaultViewName:** The name of the default view within the workbook.
4. **defaultViewRepoUrl:** The repository URL of the default view within the workbook.
5. **numberOfFavorites:** The number of times the workbook has been marked as a favorite by users.
6. **title:** The title of the workbook.
7. **viewCount:** The number of times the workbook has been viewed.

**workbookRepoUrl:** The repository URL of the workbook.

## INITIAL PLAN

My initial plan was to use Streamlit—a tool for building and deploying data apps—to fetch responses from the Tableau Public API and display them in a simple interface. For example the below looking at Prasann's journey.

### Prasann Prem's Tableau Public Journey

Using Tableau Public since: 4 years 9 months, 25 days  
 Total Number of Workbooks Created: 34  
 Total Number of Followers: 838  
 Total Number of Following: 419

### Top Workbooks by Views:

The Vedic Personality Test- Think like a Monk by Jay Shetty- 92155 views  
 Sales Analysis Dashboard #VOTD - 48087 views  
 The Color of Sleep #VOTD - 23780 views

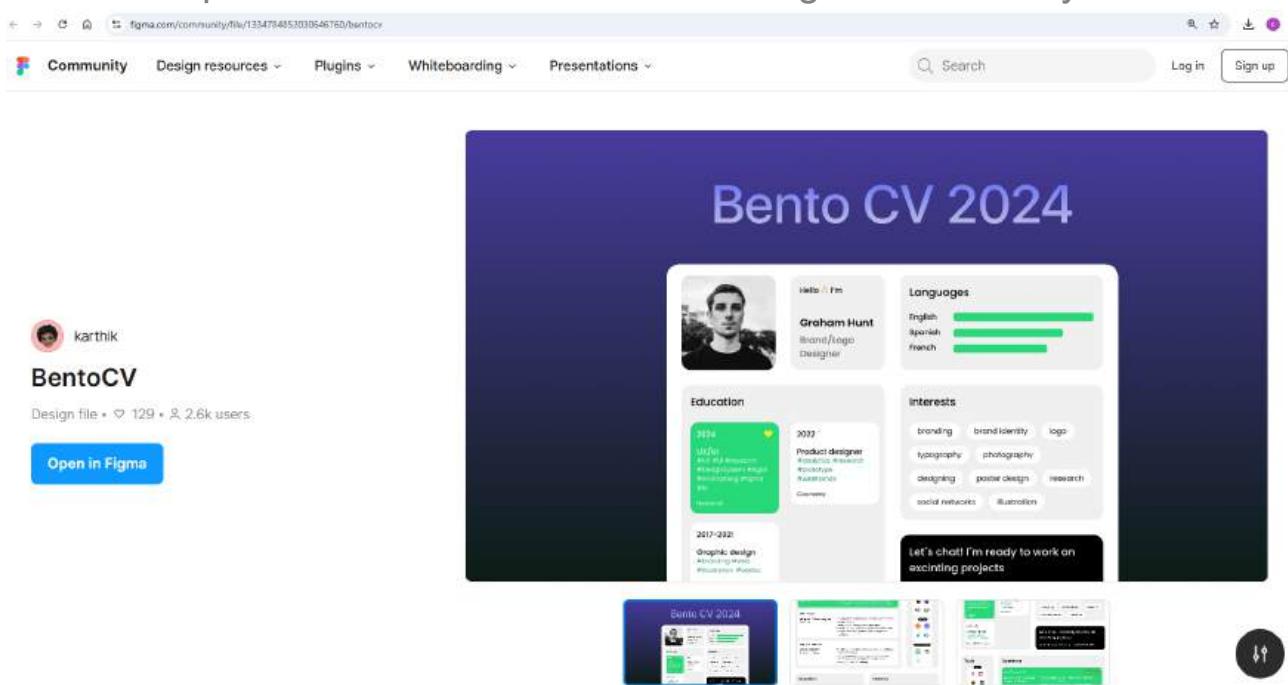


The first iteration of the Streamlit app gave me a starting point to think about the UI and how I wanted to structure the final output. One of my primary goals was to enable users to download and proudly share their achievements on social media.

However, after a little bit of researching and exploring, I found that building a simple web app using a React framework will be a better option for the project considering the long term implementations I had in mind.

## UI DESIGN

While exploring the Figma Community designs, I came across a BentoCV design by **Karthik**. To save time and jump into development quickly, I decided to customize this CV design into a Wrapped design. I then spent the next 2–3 hours fine-tuning it to match my vision.



Throughout the process, my primary focus was simple: to create something I'd personally be excited to use. If I wouldn't want to use it, I couldn't expect others to either.

The final UI design looked something like this

## Edward's Tableau Public Journey

+tableau public



Hello,

**Edward  
Steven Paul**

Data Storyteller

You're publishing Tableau Public Workbooks since,

**4 years, 9 months and 4 days**

It's a journey of perseverance



You've published  
**77**  
Tableau Workbooks



You're an inspiration to  
**459**  
people who follow you



You follow  
**23**  
people who inspire you



You've got  
**456,652**  
views on your workbooks



Your workbooks have  
**1,256**  
favorites from users

### Tableau Wrapped

Created by: Hari Jeyan

Generated on: 12/21/2024

## CODING WITH CURSOR

This was my first time using Cursor AI, and I'm thrilled with how well it worked. From deciding the tech stack, installing the necessary programming languages, and setting up the environment to organizing the project structure and coding, Cursor AI proved to be incredibly intuitive. It streamlined the entire process, allowing me to implement code through simple clicks. The key was crafting clear, detailed prompts that explained exactly what I intended to achieve.

Here's the first prompt I used in Cursor AI:

Prompt:

"I'm building a Tableau Wrapped web app, which:

- Takes a Tableau Public username as input.
- Uses the Tableau Public API to fetch user statistics, including:
  - Profile info
  - Total workbooks
  - Followers/Following counts
  - Total views
  - Join date

- Displays results in a Bento grid card UI (design already created in Figma).
- Allows downloading results as an image for social media sharing.

I have 0 coding experience, so I need step-by-step guidance to implement this project.

API Documentation is available, and we have access to endpoints such as:

- Profile: <https://public.tableau.com/profile/api/{username}>
- Workbooks: <https://public.tableau.com/public/apis/workbooks?profileName={username}&start=0&count=50>

Using this prompt, Cursor AI guided me through the coding process seamlessly, making it possible to build a functional and visually appealing web app.

The screenshot shows the Cursor AI development environment. On the left, the file explorer lists components like 'BentoCard', 'Bentocard', 'BentocardGrid', 'BentocardTable', and 'BentocardTableGrid'. The code editor displays a component named 'BentocardTable' with TypeScript code. The code includes imports for React, TableauStats, and various utility functions. It defines a 'BentocardTableProps' interface and implements a 'formatNumber' function to format numbers. It also includes logic for handling 'joinNote' and 'trackDownload' functions. The terminal at the bottom shows the command 'npm run dev' being run. A sidebar on the right provides project navigation and documentation.

```

// File: components/BentocardTable.tsx
import { useState, useEffect } from 'react';
import { TableauStats } from '@types';
import { Tableau, TableauPage, TableauPageCard } from 'react-tableau';
import HeadlessTable from 'headless-react';

interface BentocardTableProps {
  stats: TableauStats;
}

const formatNumber = (num: number) => {
  return new Intl.NumberFormat('en-US').format(num);
};

const getJoinNote = (joinNote: number | null) => {
  console.log(`joinNote value: ${joinNote}`);
  if (joinNote === null || joinNote === undefined || joinNote === 0) {
    return {
      badge: 'Legacy Creator',
      description: 'Part of Tableau's history since the early days',
      class: 'Legacy-member'
    };
  }
  return null;
};

const trackDownload = async (username: string) => {
  try {
    // Only catch in production
    if (process.env.NODE_ENV === 'production') {
      await fetch('/api/track', {
        method: 'POST',
        headers: {
          'Content-Type': 'application/json',
        }
      });
    }
  } catch (err) {
    console.error(err);
  }
};

const BentocardTable = ({ stats }: BentocardTableProps) => {
  const [joinNote, setJoinNote] = useState(null);

  const handleDownload = () => {
    trackDownload();
  };

  const handleJoinNote = (e: React.ChangeEvent<HTMLSelectElement>) => {
    setJoinNote(e.target.value);
  };

  const BentocardTableGrid = () => {
    return (
      <Tableau
        stats={stats}
        joinNote={joinNote}
        onDownload={handleDownload}
        onJoinNote={handleJoinNote}
      />
    );
  };

  return (
    <TableauGrid
      title="Tableau Grid"
      description="View your Tableau Public grid in a clean, modern design with a Bento-style grid card UI."/>
  );
};

export default BentocardTable;

```

Once I provided Cursor AI with the context of the project and set up the necessary components, the process became much smoother. All I needed to do was download the Bento UI as a PNG image from Figma, upload it to the Cursor AI chat, and instruct it to build the frontend UI. Of course, not everything worked perfectly on the first try. There were moments when the output didn't meet my expectations. In those cases, I simply refined my prompts, providing more detailed context and clearer instructions. The key was persistence—continually pushing Cursor AI to understand my requirements better.

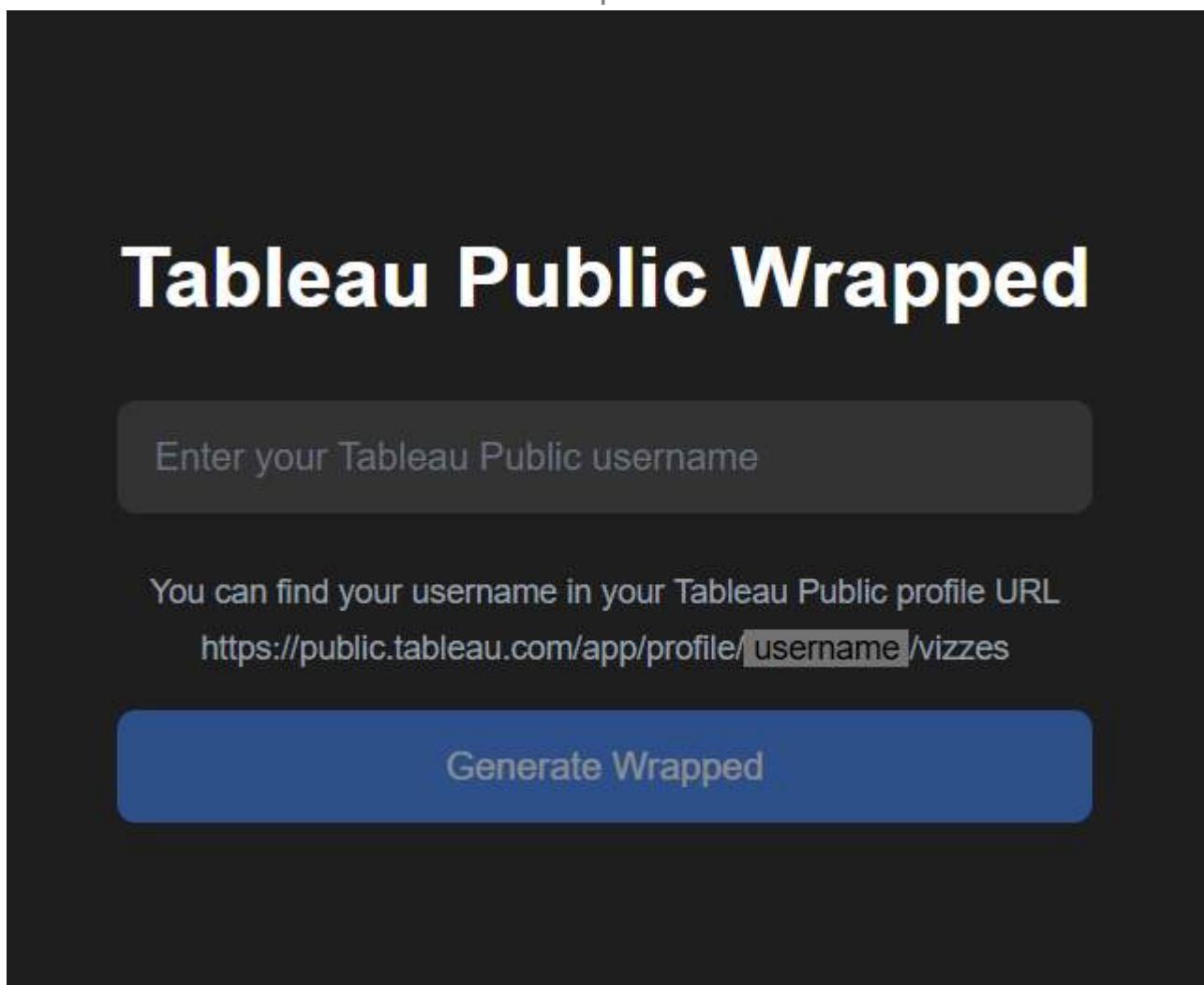
## BUG FIXES

This was the one area where I faced some challenges. If you have a little bit of experience in coding and debugging, it will be easy for you to debug as you would know what to look for.

When fixing bugs, it's essential to provide clear context about the issue and add the specific code file to the Cursor AI chat. This approach not only saves time but also prevents unintended changes to the working parts of the code.

### FINALISING THE PROJECT

I have nearly checked 100+ times with various user names, fixed small things to make it look better in each iteration, and made it mobile responsive, as I expected more people would access this from their mobile phones.



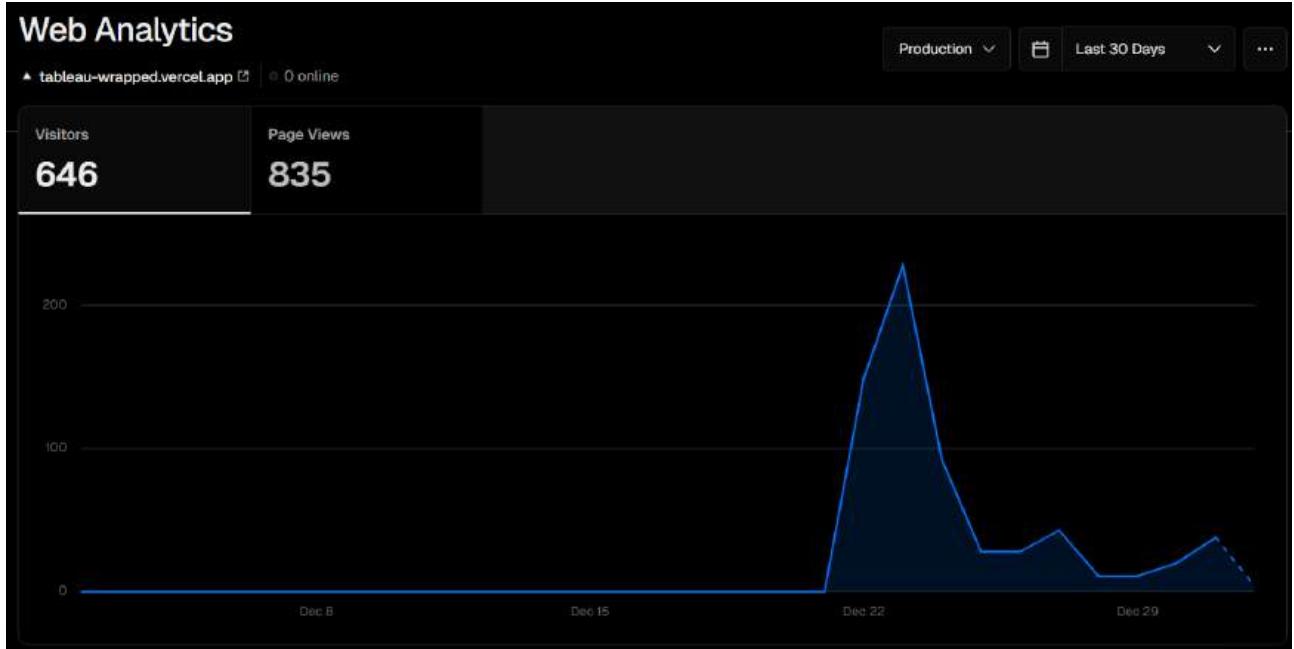
It felt like reinventing myself after some time. Doing things for the first time and getting it done perfectly before the party is over was the best feeling I had in a while.

## SHARING WITH THE COMMUNITY

Initially, my plan was just to make a simple LinkedIn post and share the link in the post. I thought there was nothing much to say about it as it is a simple web app that gives the Tableau Public Wrapped when a user enters their Tableau Public user name, which can be found in their profile URL.

But when I shared my plan with my wife, who is a Machine Learning Engineer and not a Tableau user, she suggested that it would be better to make a short video on how to use this web app and what it does without writing any lengthy post and people would stop scrolling and check what this is about. But I was scratching my head and thinking, what's the point of doing all this? The only thing I didn't understand was the perspective. Just because I spent two full days building this and know the ins and outs of the product and how it works, for someone who's seeing this for the first time may not see this product in the same way as I see it. To overcome this challenge a short explainer video would be the best thing.

I recorded my screen, found a promo video template in Canva, and quickly edited the Tableau Promo video, which went well beyond my expectations. I shared this on a Sunday evening. The responses I received were overwhelming, and I was so glad to see people try this app and share their Tableau Public Wrapped with the pride of their Tableau journey.



## REFLECTIONS

If you'd asked me two years ago whether I could have accomplished this project, my answer would likely have been a "no." Even if I had the idea back then, I wouldn't have been willing to relearn coding or dedicate weeks to building something like this. It would've taken weeks, and the party would've been over.

However, with the arrival of Generative AI, the landscape has changed.

What once seemed impossible is now just challenging—something achievable with commitment and the right tools. Gen AI has lowered the barriers to entry, making ideas that once seemed out of reach more attainable.

## NETWORKING

Before this project, I followed many top 1% contributors in the Tableau datafam, but I wasn't directly connected to most of them. After sharing my project, things changed. I received DMs, gained followers, had my post shared and tagged by others, and built meaningful connections.

This blog also happened because of that.

This experience made me realize the truth in Naval Ravikant's words: **"Networking is overrated... Do something great, and your network will instantly emerge."**

## FUTURE PLANS

There are a few improvements I'm planning to add to this Tableau Wrapped app with time.

- Currently, the downloaded Tableau Wrapped image on desktop and mobile looks different based on responsiveness. Make the downloaded image look the same across all devices.
- The individual dashboards have created dates and last updated dates. Using this, make a dashboard that visualizes annual user engagement trends in Tableau Public.
- Enable the user to set annual, quarterly, or monthly goals for Tableau Public and measure the streak in GitHub style.

I hope this was worth the read. Along with the project's technical details, I wanted to inspire people that they could do things with the help of AI that seemed impossible two years ago.

### Hari's Tableau Public Journey





Hello,  
**Hari Jeyan**  
Data Analytics Engineer

You're publishing Workbooks since,  
**2 years, 8 months and 7 days**  
It's a journey of perseverance

 You've published  
**24**  
Tableau Workbooks

 You're an inspiration to  
**9**  
people who follow you

 You follow  
**74**  
people who inspire you

 You've accumulated  
**1,561**  
views on your workbooks

 Your workbooks have  
**1**  
favorites from users

**#TableauWrapped**  
Created by: Hari Jeyan  
Generated on: 02/01/2025

X  
LinkedIn  
Tableau Public  
GitHub

CJ Round Up: Thank you so much for sharing Hari the inner workings of what you have created! I love that it shares everything from design, to BI tools, to code, and suggestions along that journey.

LOGGING OFF,

CJ