Unveiling Insights:
## The Power of Data Visualization with Tableau

By
Shri Abhiraami Thangavel
Jan 2024

A little about me!

———

- Architect and Project Manager turned Data Scientist
- Recent graduated from Michigan State University in Business Data Science
- Currently interning at Roosevelt Innovations
- Always had the love for visual storytelling
- Short term Goal: PowerBI Certification - PL300 and Tableau Desktop Specialist Certification

NARRATIVE

VISUALS

ENGAGE

EXPLAIN

ENLIGHTEN

SWEET SPOT
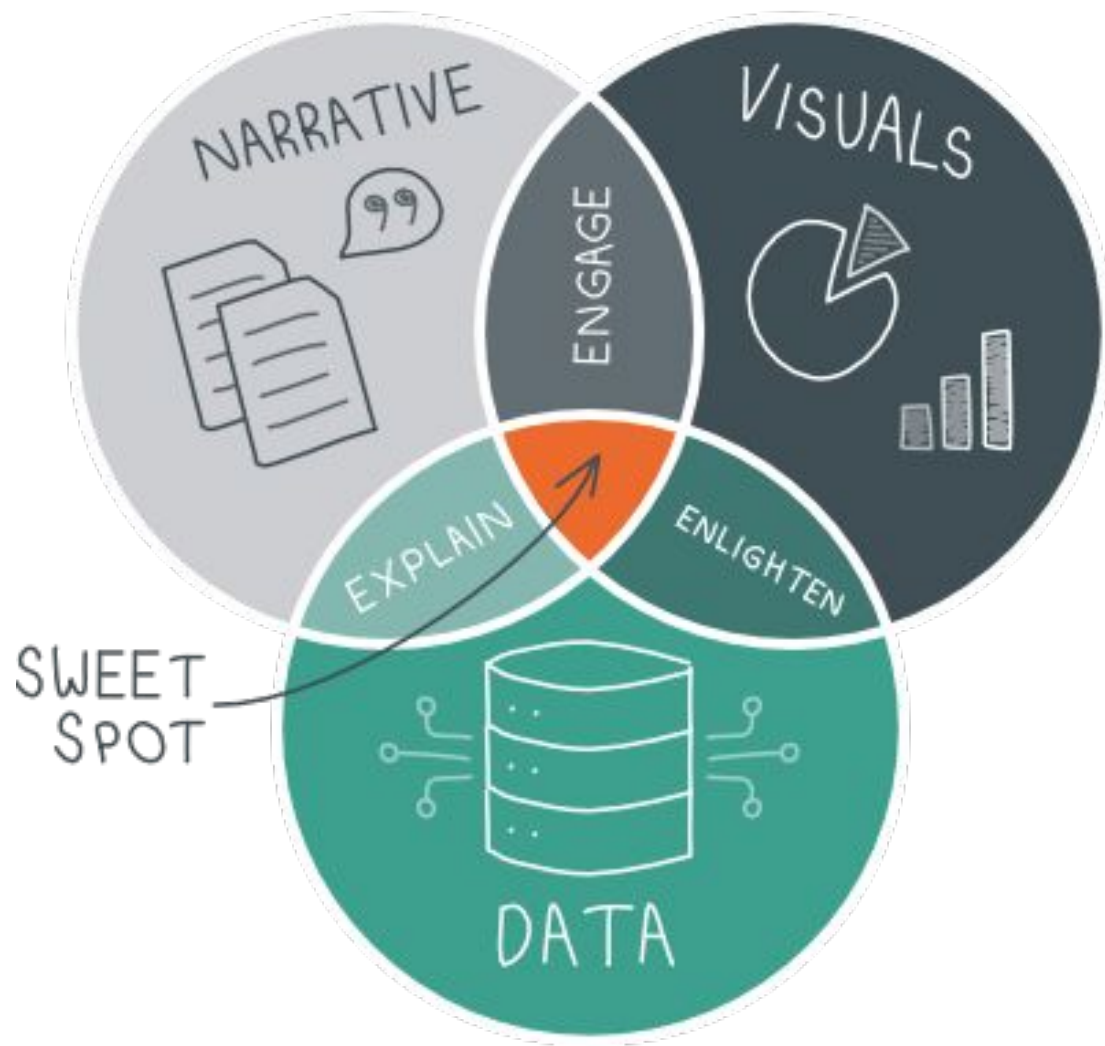
DATA

What is Data Visualization?

———

"The representation of data in a visual format to gain insights and communicate findings effectively."

Importance:
1. Clarity:
        Visualizations simplify complex data for easier comprehension
2. Insight Discovery:
        Uncover patterns, trends, and outliers
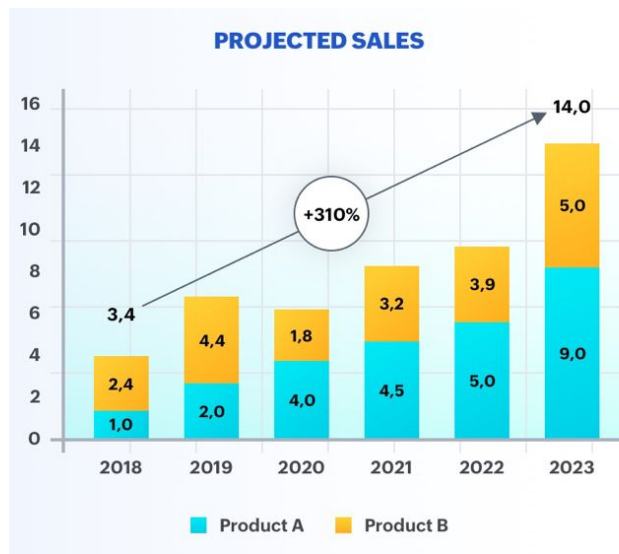3. Decision-Making:
        Impact of clear visualizations on informed decision-making

Why do we have to visualize data?

___

Allows audiences to grasp key insights and trends at a glance

General Flow for Data Visualization

Data Check → Explore Data → Analyse and Visualize Data → Dashboarding → Communicate Insights
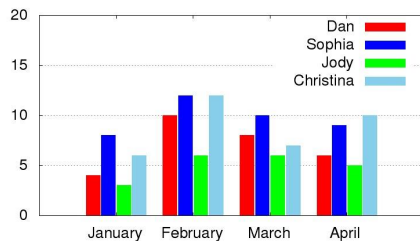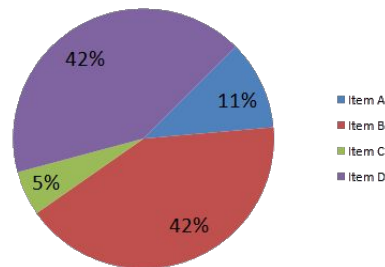
Different Visualization Types

— — —

PROJECTED SALES

| Year	Product A	Product B
2018	1,0	2,4
2019	2,0	4,4
2020	4,0	1,8
2021	4,5	3,2
2022	5,0	3,9
2023	9,0	5,0

Tables: Consist of rows and columns and are used to compare variables in a structured way.

Bar charts: Use vertical or horizontal bars to compare categorical data. Used for analyzing value trends.
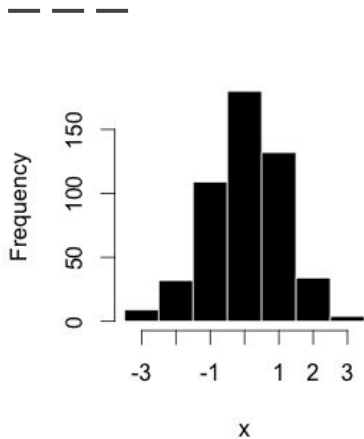

Pie charts: Used to compare the size of each component and are usually used to determine a percentage of the whole.
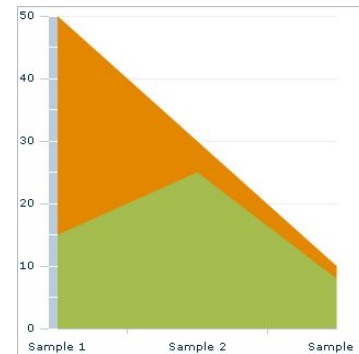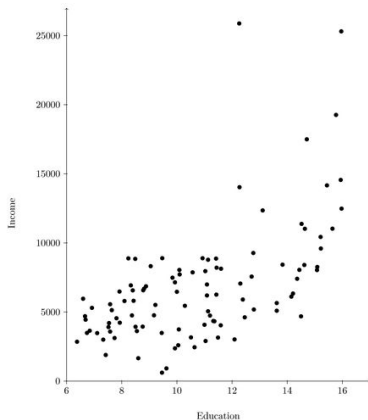

Area charts: Similar to bar and line graphs and show the progress of values over a period. Showcase data with a time-series relationship

Different Visualization Types

Histograms: Distribute numerical data. They are mainly used to plot the distribution of numbers and analyze the largest frequencies within a particular range.

Scatter plots: Present the relationship between two variables. Show trends, clusters, patterns, and outliers in large dataset.

Heat maps: Values in a matrix are depicted using colors to convey the intensity of the data, with warmer colors typically indicating higher values and cooler colors indicating lower values.
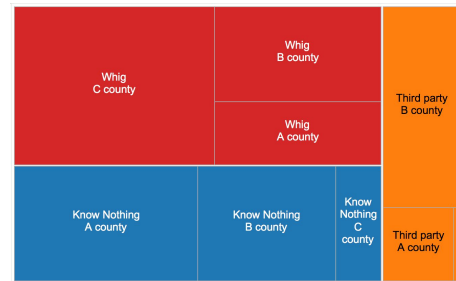
Tree maps: Hierarchical data in nested rectangles, with each ponds to a quantitative value.

Introduction to Tableau

———

"Tableau is a powerful data visualization tool that allows users to create interactive and shareable dashboards."

Advantages:
1. User-Friendly Interface:
        Intuitive design for easy usage.
2. Versatility:
        Adaptable to various data types and
   sources.
3. Interactivity:
        Showcase interactive features for
   enhanced data exploration.

Excel vs Tableau

— — —

	Excel	**Tableau**
Advantages	Familiar Interface, widely used. Handles basic data analysis tasks. Available to most users.	Creates dynamic and interactive visuals. Seamless integration with various sources. Real-time sharing and collaboration features.
Limitations	Challenges with large datasets. Limited capabilities for advanced visualizations. Mostly static charts and graphs.	Steeper learning curve Separate license and may have associated costs Relies heavily on online connectivity for certain features.
Use Cases	Simple calculations and basic analysis. Tabular data organization.	Complex visualizations. Interactive dashboards for diverse audiences.

The Problem

———

A fictitious dataset about churn from a Telecom Provider

Our Task: DISCOVER WHY CUSTOMERS ARE CHURNING

What is churn/churn rate?
The rate at which customers stop doing business with an entity.
Basically: **Churn Rate: Customers Lost/ Total Customers**
    Leaky Bucket Problem
    Getting new customers is difficult than keeping existing customers

About the Data

———

Key Characteristics
    Contains 29 columns
    One row per customer, total rows: 6688

Dimensions

Column name	Description
Customer_id	The unique ID that identifies a customer
Churn Label	Contains 'Yes' or 'No' to indicate if a customer churned.
Demographic fields	Age, Gender, State, ...
Premium plans	Unlimited Data, International Plan, ...
...	...

Measures

Column name	Description
Total charges	Sum of all monthly charges
Monthly charge	The average of all monthly charges billed to the customer
Extra data charges	Extra charges for data downloads above the specified customer plan
Extra international charges	Extra charges for international calls for customers not on an international plan
Customer service calls	Number of calls made to customer service
...	...

Let's start analyzing!

Best Practices for Effective Data Visualization

———

1. Simplicity:
   Keep visuals straightforward to enhance audience understanding.

2. Consistency:
   Maintain a uniform style throughout dashboards using fonts & colours.
   Enhances overall aesthetic and user experience.

3. Storytelling:
   Utilize Tableau to tell a cohesive and compelling data-driven story.
   Capture audience attention and convey insights in a narrative
format.

It's your turn!

Dataset

STEP 1: Download the Titanic Dataset from Kaggle

(Link: https://www.kaggle.com/c/titanic/data)

STEP 2: Open the Train.csv on excel.

Let's try to answer these questions

___

1. No of Rows and Columns??   Rows: 891 Columns: 12

2. Seeing the column labels,   Dimensions: Survived, Sex,
   what are the dimensions   Age, Embarked, PClass
   and measures?
                                     Measures: Fare

Dataset

———

STEP 1: Download the Titanic Dataset from Kaggle

(Link: https://www.kaggle.com/c/titanic/data)

STEP 2: Open the Train.csv on excel.

STEP 3: Load the Train.csv on Tableau

Let's try to answer these questions

———

3. How many survived?   342

4. No of Male and Female who   Male: 577, Female: 314
boarded the ship?

5. Which gender survived the   Female
most?

Let's try to answer these questions

———

6. In which class, did most people travel in?

Class 3

7. Which class survived the most?

Class 1

(Bonus: Check the survival across Class based on Gender)

8. Which age group of travellers is the highest?

20-25

Let's try to answer these questions

———

9. From which port had the least number of Embarkation?

C = Cherbourg, Q = Queenstown, S = Southampton

10. Give a overall narrative or communicate your insights which these visualizations we created!

Q-Queenstown

1. Higher Class Female Survived the most
2. Lowest Class Male were given the least preference.

Funnily, that what we saw in the movie!

Thank you!

Q & A