Build Your Online Portfolio

Data Science Bootcamp

Contents

- Review data science bootcamp
- Before you apply for a job
- SEO = discovery
- Must-have portfolios
- Let’s create your own website
- Let’s create your own blog (medium)

Bootcamp Review

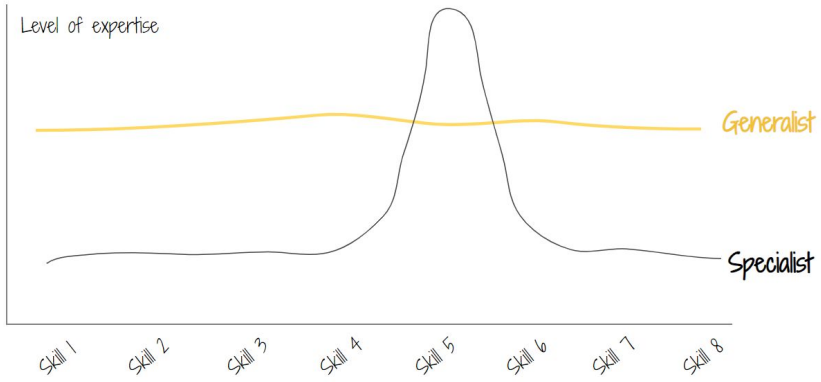
Let's do survey

https://forms.gle/JujUHZD4kwqbd3r68

You are now full-stack Generalist

- Spreadsheets
- SQL
- Data Transformation | Visualization
- Programming skills in R, Python
- Statistics | ML Basics
- Dashboard
- + Other useful knowledges

When great becomes enemy of good



Diminishing Return

Skills in combination is more powerful than an individual skill

Skill Checklists

Spreadsheets
- Functions
  - IMPORTDATA
  - IF
  - FILTER
  - SORT
  - VLOOKUP
  - XLOOKUP
  - QUERY
  - REGEX_
  - CORREL
  - LINEST
- Pivot Table
- Charts
- Power Query
- Analysis Toolpak
- Solver (Optimization)

SQL
- SELECT
- WHERE
- JOIN Two Tables
- JOIN Multi Tables
- Aggregate Functions
- Subqueries
- Window Functions
- CREATE TABLE
- UPDATE TABLE
- Using SQL with R, Python

Python
- Google Colab
- DataLore Notebook
- Variable
- Data Types
- Data Structures
- Functions
- Classes/ OOP
- Pandas
- Requests/ API Basics

Dashboard
- Google Data Studio
- Power BI
- Tableau Public
- Connect Data
- Data Model
- Dimension/ Metrics

R
- Programming Basics
- Data Transformation
- Data Visualization
- Report using markdown
- Working with DataFrame
- Connect to SQL
- Analyze data basics

Useful Data Knowledge
- Google Analytics
- SEO
- Basic Marketing

Statistics
- Descriptive
- Point Estimate
- Interval Estimate
- Correlation
- Linear Regression
- Logistic Regression
- AB Testing
- P-value

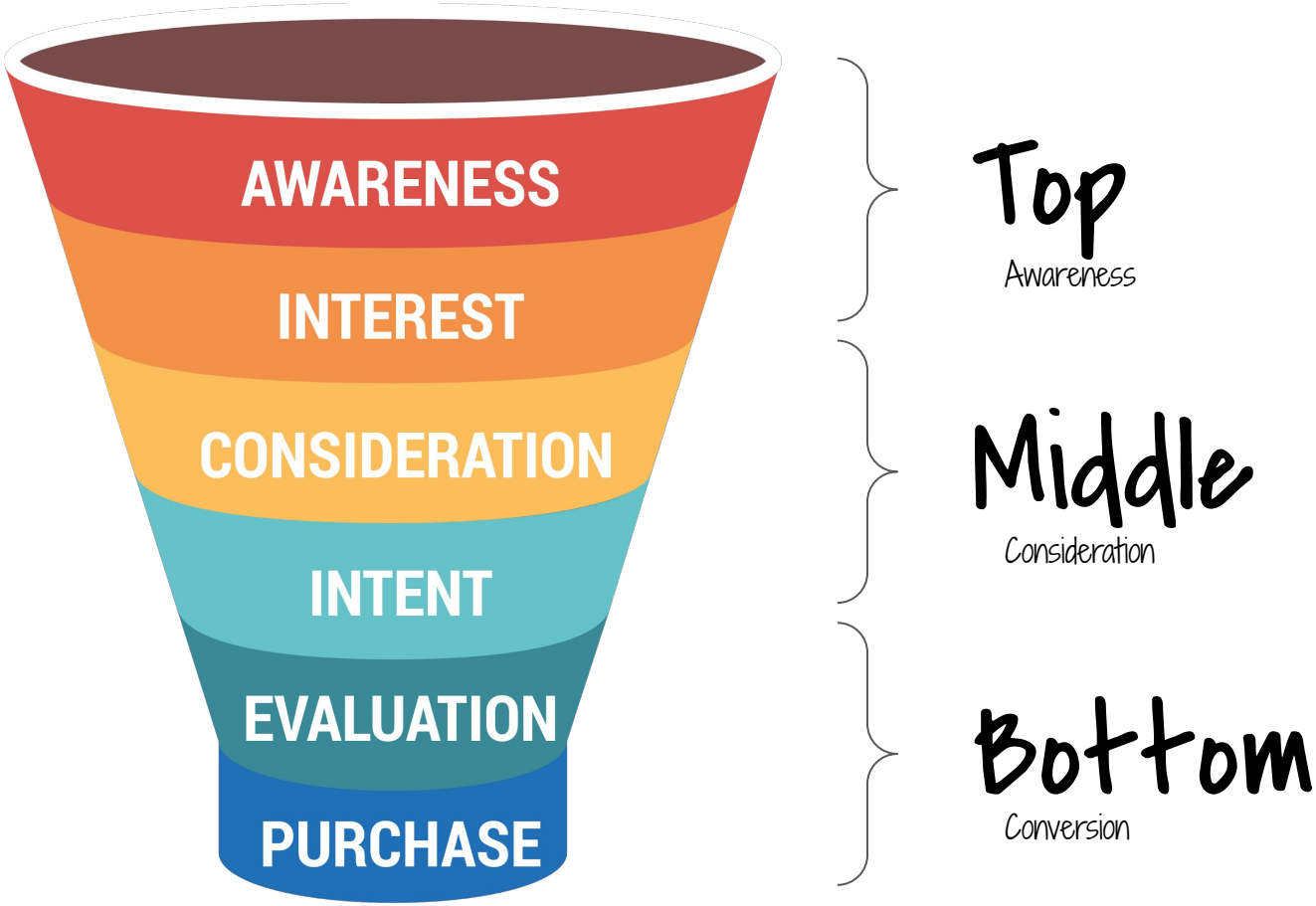
Machine Learning
- Basic Models
- Train Test Split
- K-Fold
- Bootstrap
- Leave One Out CV
- Regression
- Decision Tree
- Random Forest
- K-NN model
- K-Means Clustering

Core Business Metrics

Marketing Funnels

Business Domain

- Economics : demand/ supply
- Business Metrics
- Digital Marketing
- Google Analytics
- SEO
- Consumer Behaviours


AWARENESS
INTEREST
CONSIDERATION
INTENT
EVALUATION
PURCHASE
Top
Awareness
Middle
Consideration
Bottom
Conversion

Awareness

- Brand awareness
- Reach

Consideration

- Traffic
- Engagement
- App installs
- Video views
- Lead generation
- Messages

Conversion

- Conversions
- Catalog sales
- Store traffic

Pirate Metrics

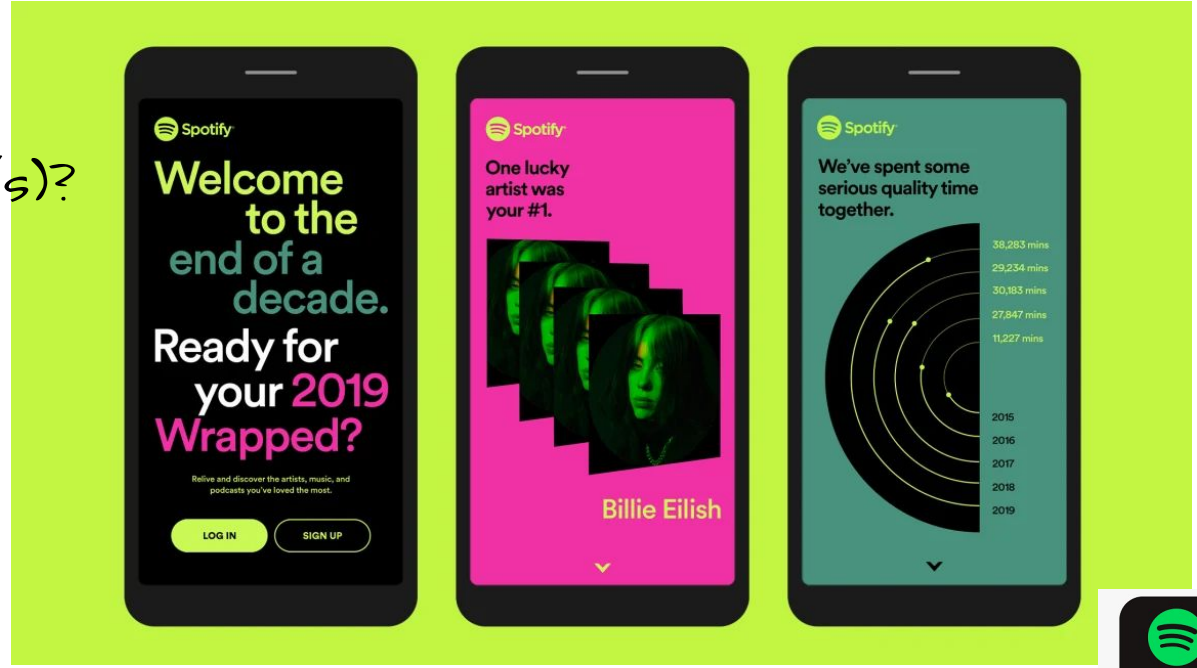
AARRR

1. Awareness
2. Acquisition
3. Retention
4. Referral
5. Revenue

Case 1

Improve which metric(s)?

1. Awareness
2. Acquisition
3. Retention
4. Referral
5. Revenue



Case 2 churn prediction model

Improve which metric(s)?

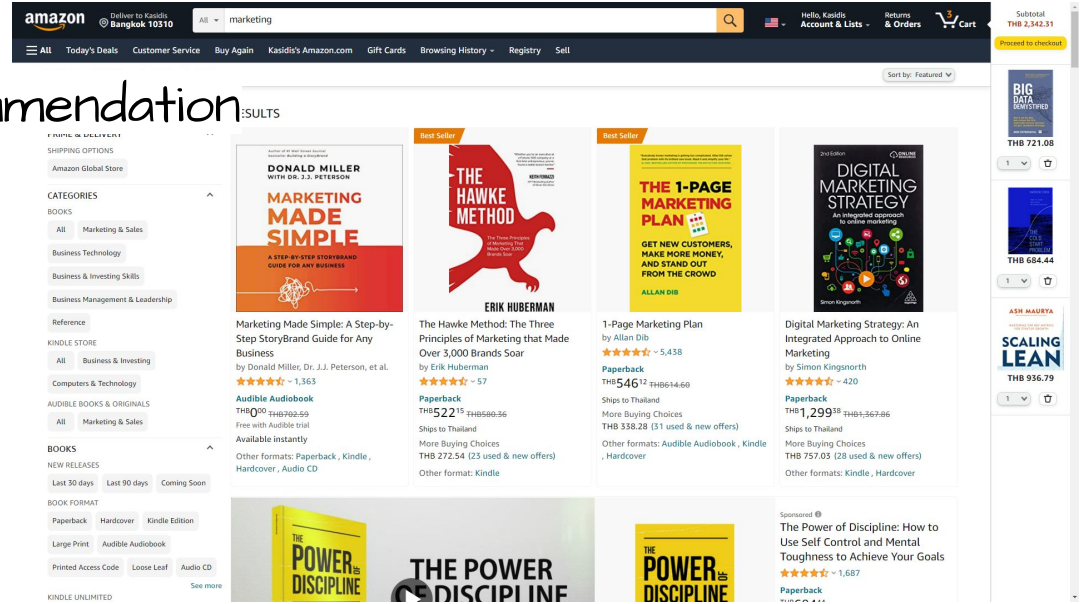
1. Awareness
2. Acquisition
3. Retention
4. Referral
5. Revenue

amazon

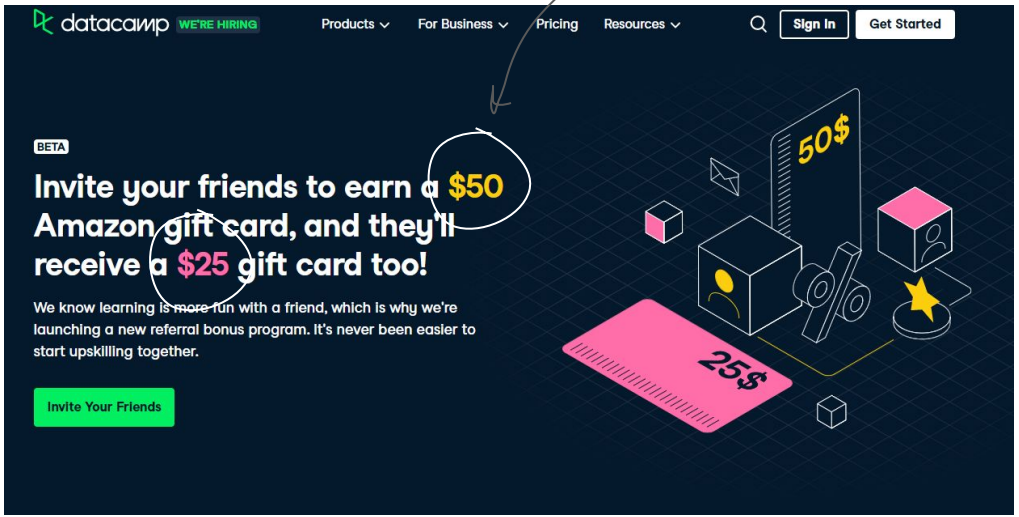
Case 3 product recommendation

Improve which metric(s)?

1. Awareness
2. Acquisition
3. Retention
4. Referral
5. Revenue



use data to figure out this number



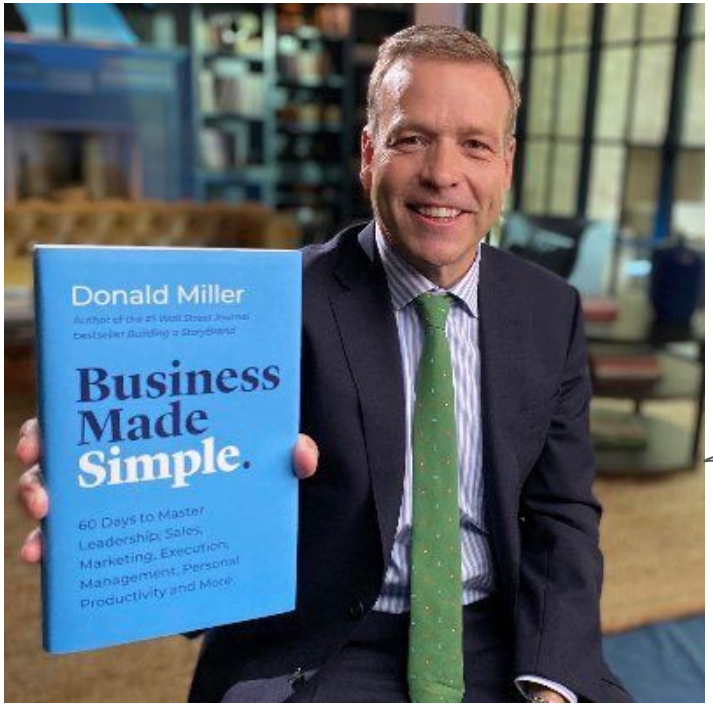
Case 4 referral program

Improve which metric(s)?

1. Awareness
2. Acquisition
3. Retention
4. Referral
5. Revenue

Business Made Simple

Understanding business in 10 minutes



Donald Miller

Business Made Simple

Be a value driven professional

Economic Product

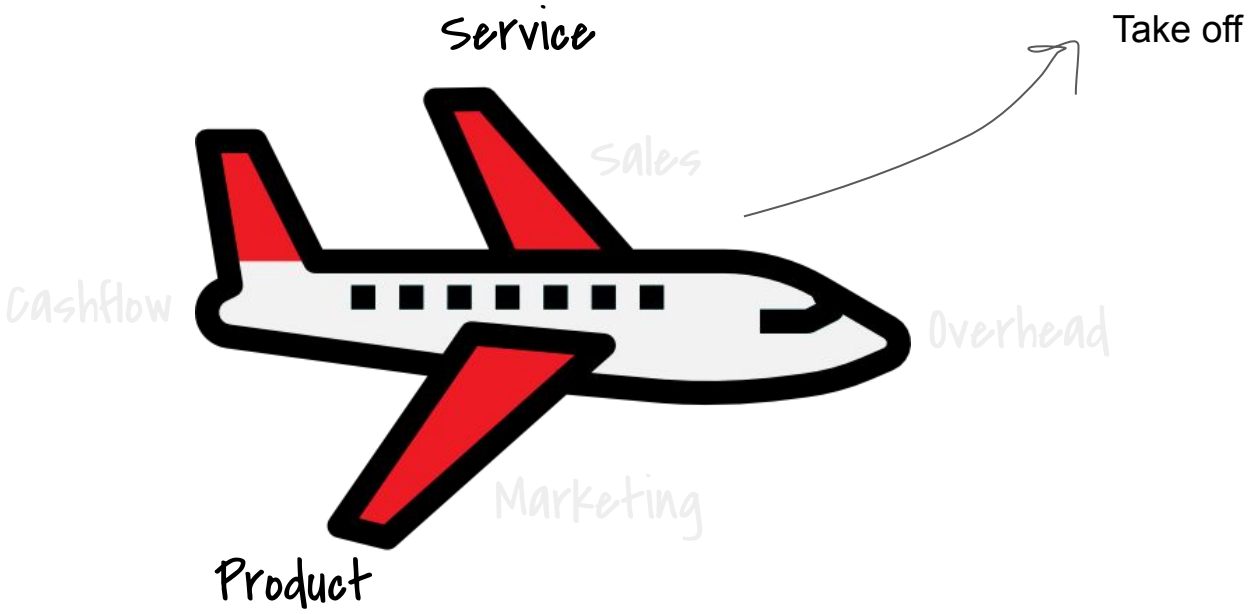
Business

Service
Sales
Cashflow
Overhead
Marketing
Product

Body

Overhead : rent, salary, office supplies, gas billing

Wings


Service
Take off
Sales
Cashflow
Overhead
Marketing
Product

Engines

Fuel

Service

Sales

Cashflow

Overhead

Marketing

Product

> Nothing sends up a red flag for your boss faster than a conversation in which it becomes obvious that **you don’t understand cash flow**.
>
> Miller, Donald. Business Made Simple (p. 91). HarperCollins Leadership. Kindle Edition.

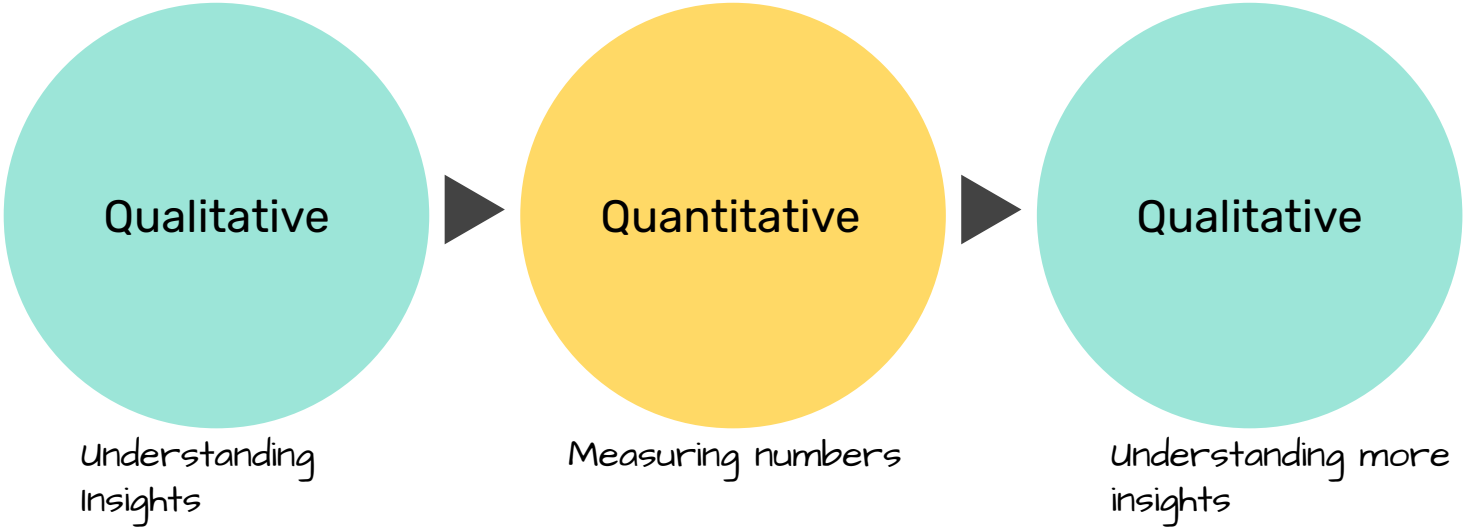
Balance

Research Skills

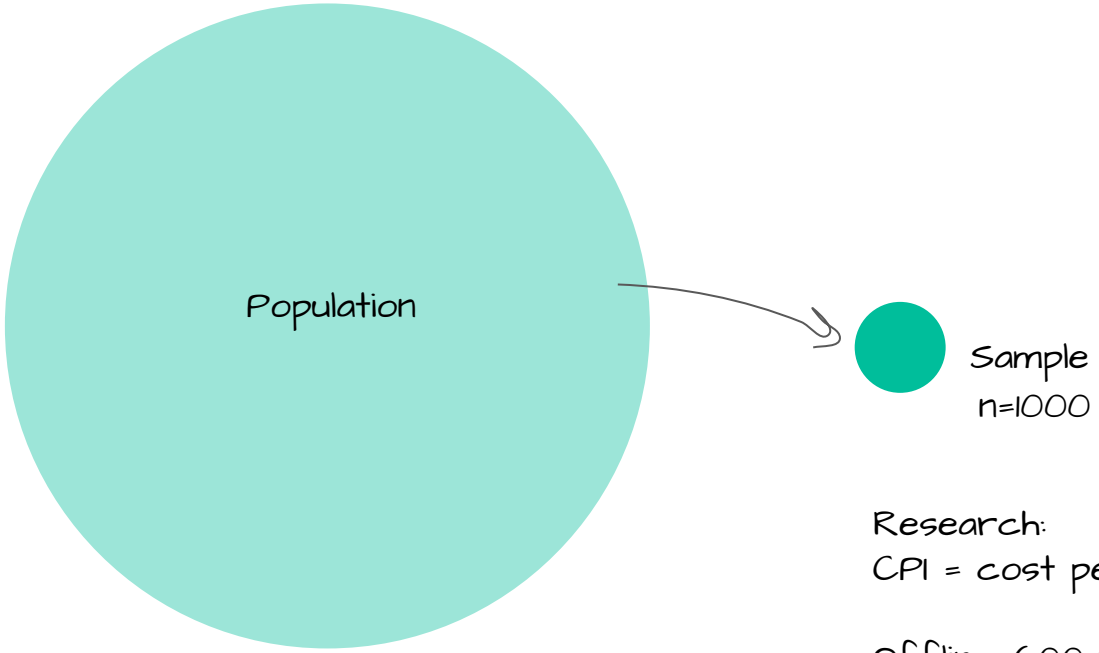
Marketing Research

Market Research

- Quantitative
  - Home Survey
  - Online Survey
  - Street Intercept
  - Panel
- Qualitative
  - Focus Group
  - In-depth Interviews
  - Ethnography


Qualitative
Quantitative
Qualitative
Understanding Insights
Measuring numbers
Understanding more insights

Big data is not enough (What, How)
Small data can answer the **WHY**



Research:
CPI = cost per interview

Offline: 600-800 THB per interview
Online:

Point Estimate

Confidence

60% [55%, 65%] at 95% confidence level

Interval Estimate

What You Need

Before applying for a job

Must Have!

- LinkedIn
- Resume (PDF) using google doc/ canva
- Online portfolio/ personal website

\+ Certifications (from well-known institutions)

Coursera highly recommended

SEO

make yourself more discoverable

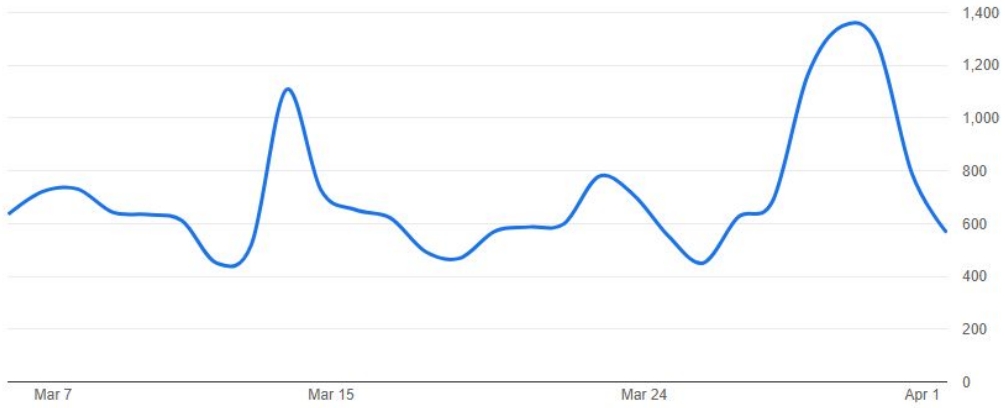
Search Engine Optimization

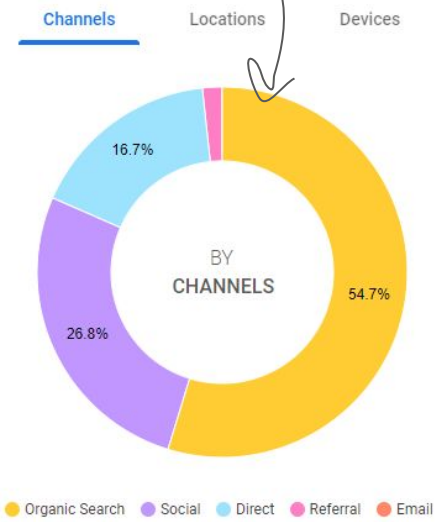
DataRockie
SEO is free

All Users

16K

↑ 24% compared to the previous 28 days





Source: Analytics

Google

คอร์สเรียน data science

All | Images | News | Videos | Maps | More | Tools

About 74,700 results (0.28 seconds)

https://www.designil.com › 8-data-sci... · Translate this page

แนะนำ 8 คอร์สเรียน Data science เข้าใจง่ายได้ความรู้ - Designil

รวมคอร์สคุณภาพเยี่ยมทั้งภาษาไทยและภาษาอังกฤษทางด้าน **Data Science** ที่เปิดสอนอย่างต่อเนื่อง เรียนจบแล้วสามารถทำงานได้จริง ...

คอร์สเรียน Data Science ภาษาไทย · TUXSA · คอร์สเรียน Data ภาษาอังกฤษ · Data Camp

https://www.chulamoocachieve.com › ... · Translate this page

Data Science Pathway คอร์สเรียน Data Science - CHULA ...

CHULA MOOC Achieve เปิดคอร์สเรียนออนไลน์หลักสูตร **Data Science** ศาสตร์แห่งการเก็บรวบรวม การจัดการ วิเคราะห์ข้อมูล Data Analytics และ Big Data ลงเรียน Data ...

Oct 30, 2561 BE · Uploaded by CHULA MOOC Achieve

https://victorytale.com › 8-online-dat... · Translate this page

แนะนำ 10 เว็บเรียน data science ออนไลน์สำหรับผู้สนใจทุกระดับ

อย่างไรก็ดีคอร์สเหล่านี้จะเป็นคอร์สที่สอน **data science** โดยรวม ไม่ใช่คอร์สที่เน้นทักษะด้านใดด้านหนึ่ง (...

Dec 9, 2562 BE

Udacity · Udemy · Datacamp · Coursera

https://datarockie.com · Translate this page

DataRockie | เริ่มเรียน Data Science ง่ายๆด้วยตัวเอง

แหล่งรวมความรู้ บทความ คอร์สเรียนปูพื้นฐานสำหรับงาน **Data** Analyst, Coding และ Business ฟรี สำหรับคนไทย.

url

title

https://datarockie.com · Translate this page

DataRockie | เริ่มเรียน Data Science ง่ายๆด้วยตัวเอง

แหล่งรวมความรู้ บทความ **คอร์ส**เรียนปูพื้นฐานสำหรับงาน **Data** Analyst, Coding และ Business ฟรี สำหรับคนไทย.

description

https://blog.datath.com › data-science... · Translate this page

อยากเป็น Data Scientist คุณเองก็เริ่มได้! คุยกับคนไทยที่เรียนคอร์ส ...

คอร์สออนไลน์ด้าน **Data Science** มีมากขึ้นเรื่อย ๆ จากหลากหลายมหาวิทยาลัยครับ ซึ่งมหาวิทยาลัยดัง ๆ และโปรเฟสเซอร์ดัง ๆ ก็ทำ**คอร์ส**น่าสนใจออกมาเต็มไปหมด ...

https://techsauce.co › life-hacks › fiv... · Translate this page

แนะนำ 5 คอร์สเรียนออนไลน์ Data Science ที่คุณสามารถเรียนจบได้ ...

อยากลง**คอร์ส**เรียนออนไลน์เกี่ยวกับ **Data Science** แต่ไม่มีเวลาทำยังไงดี? สำหรับคนในวงการสายเทคเป็นที่รู้กันดีว่า '**Data Science**' ...

Long form content

- Relevant to user search queries
- Useful
- Having search terms and other variants

8-10 minute read

SEO prefers long form + valuable contents

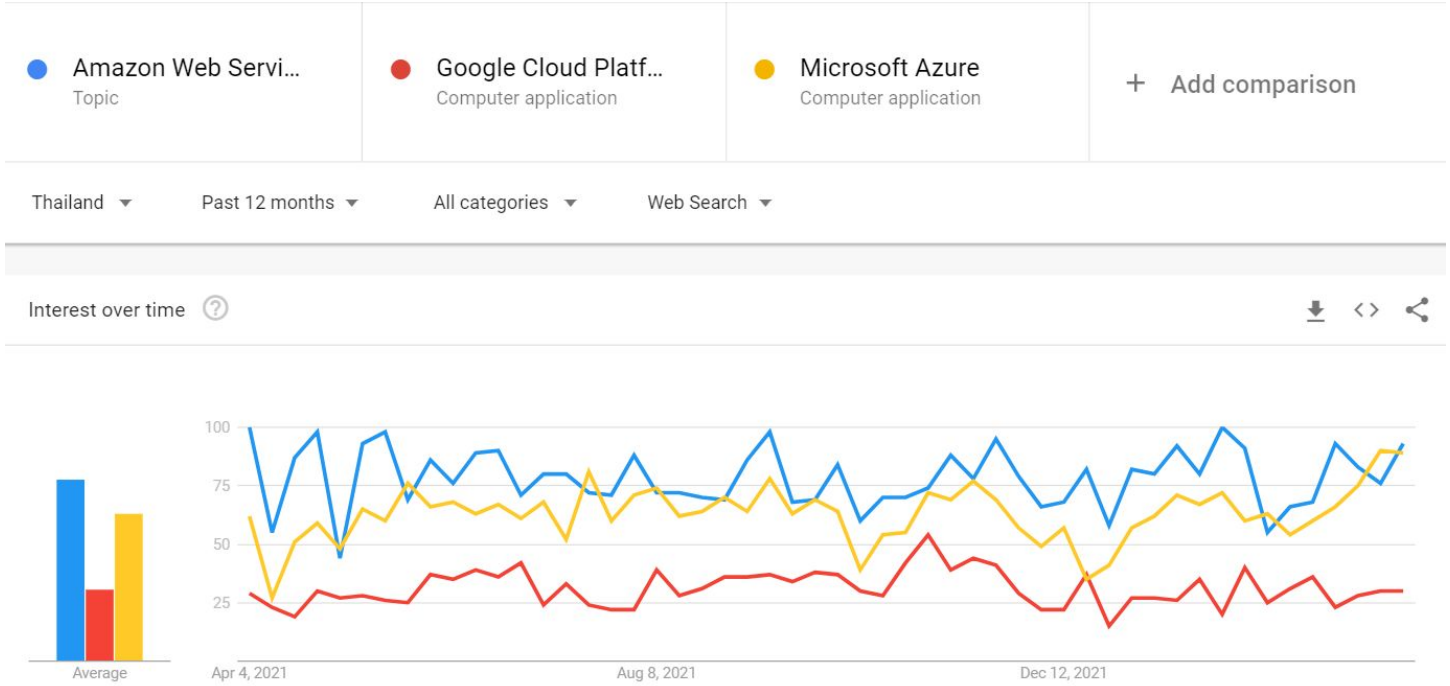
Short form content

1-2 minute read

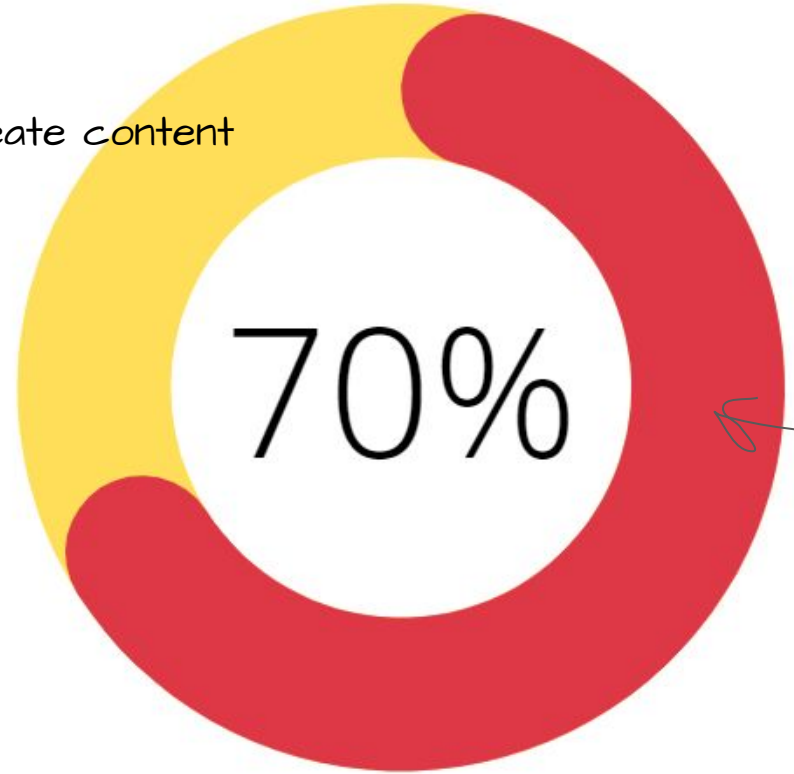
ทิปสำหรับการเขียนบทความ

- หาเรื่องที่เราสนใจ
- มีสไตล์เป็นของตัวเอง (own voice)
- time block สำหรับการเขียนบทความ
- เขียนอย่างน้อย 1-2 บทความต่อเดือน

Research new topics



https://trends.google.com/trends/explore?geo=TH&q=%2Fm%2F05nrgx,%2Fm%2F0105pbj4,%2Fm%2F04y7lrx



Promote content

find distribution channels

Tracking

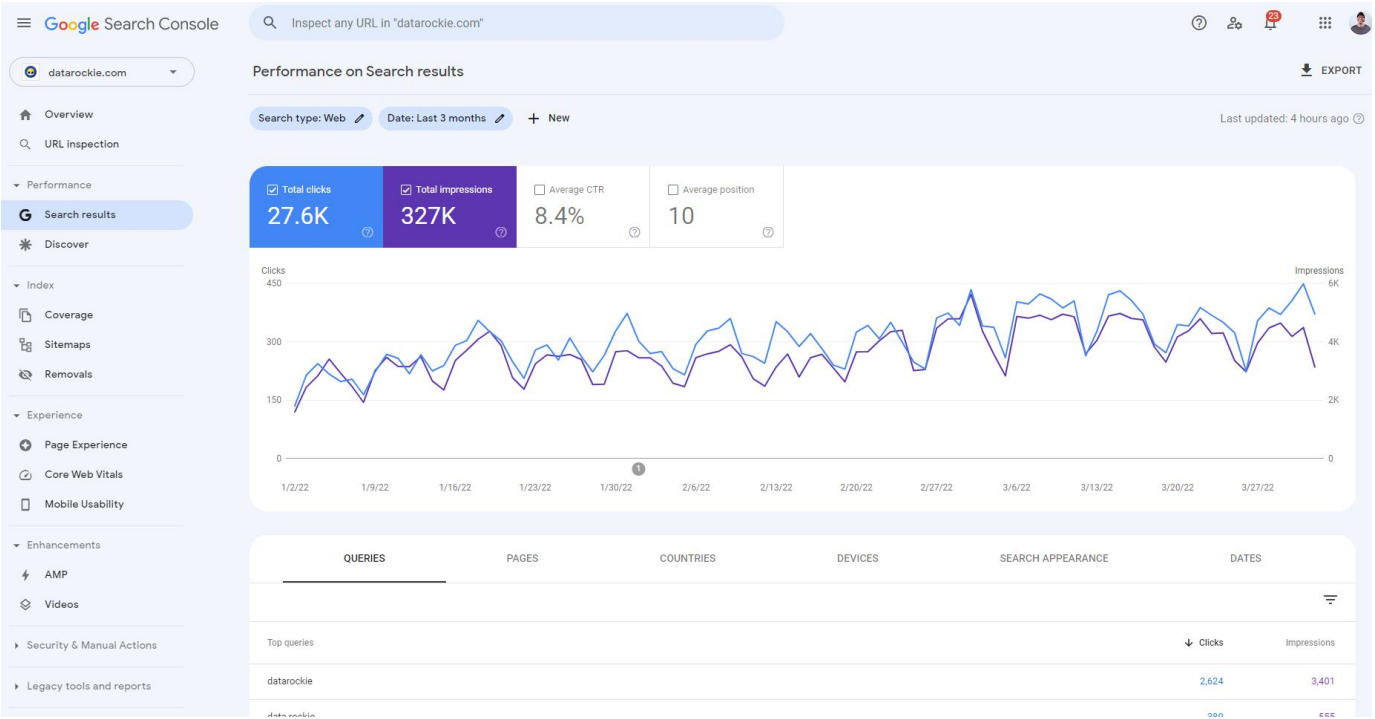
Track your website traffic

Contents

- Review data science bootcamp
- Before you apply for a job
- SEO = discovery
- Must-have portfolios
- Let’s create your own website
- Let’s create your own blog (medium)

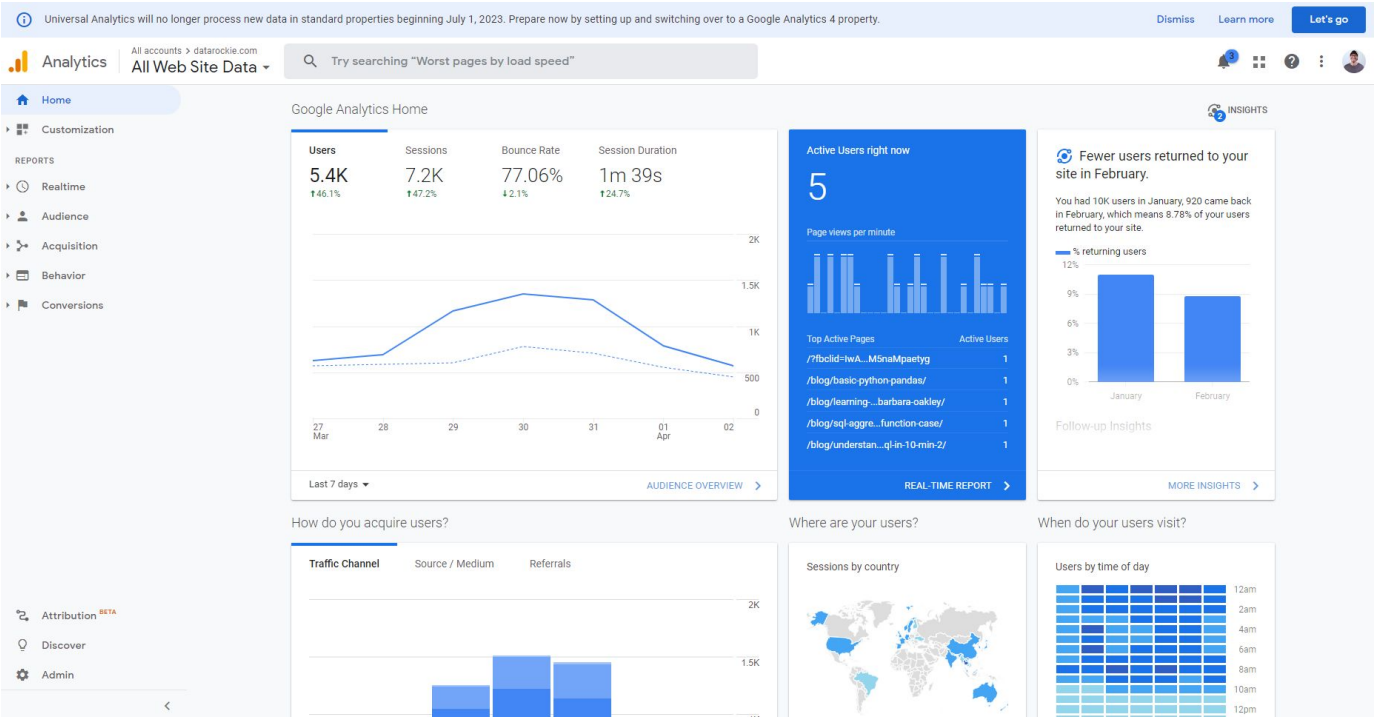
- Google Search Console
- Google Analytics

Google Search Console



https://search.google.com/search-console/performance/search-analytics?resource_id=sc-domain%3Adatarockie.com

Google Analytics



https://analytics.google.com/analytics/web/#/report-home/a59884929w279620108p247252381

Must Have Portfolios

Data, Dashboard, Model, Insights

Recommended Projects

- Spreadsheets
  - Simple dashboard / pivot table
- Databases SQL
  - Create database
  - Window functions/ subqueries / WITH
- Programming
  - R or Python
- Statistics
  - AB Test
  - Linear Regression
- Basic Machine Learning
  - 1x Regression
  - 1x Classification
  - 1x Clustering **or** 1x PCA
- Dashboard
  - Google Data Studio
  - Tableau **or** Power BI (using the same dataset) **or** both

Machine Learning Projects
1x Regression
1x Classification

Follow best practices, avoid overfitting

Extra Credit
1x Clustering or
1x PCA

Super Extra Credit
- If you can package and deploy your model
- Create API for the model

How to evaluate models using common metrics

Regression

- R-squared
- MAE
- MSE
- RMSE

Classification

- confusion matrix
- accuracy
- precision
- recall
- f1 score
- ROC (area under the curve)

Remember that in real life, performance trades off with time, budget, resources

Clustering

Observations

Filter

	Sepal.Length	Sepal.Width	Petal.Length	Petal.Width	Species
1	5.1	3.5	1.4	0.2	setosa
2	4.9	3.0	1.4	0.2	setosa
3	4.7	3.2	1.3	0.2	setosa
4	4.6	3.1	1.5	0.2	setosa
5	5.0	3.6	1.4	0.2	setosa
6	5.4	3.9	1.7	0.4	setosa
7	4.6	3.4	1.4	0.3	setosa
8	5.0	3.4	1.5	0.2	setosa
9	4.4	2.9	1.4	0.2	setosa
10	4.9	3.1	1.5	0.1	setosa
11	5.4	3.7	1.5	0.2	setosa
12	4.8	3.4	1.6	0.2	setosa
13	4.8	3.0	1.4	0.1	setosa
14	4.3	3.0	1.1	0.1	setosa
15	5.8	4.0	1.2	0.2	setosa
16	5.7	4.4	1.5	0.4	setosa
17	5.4	3.9	1.3	0.4	setosa

PCA

Features

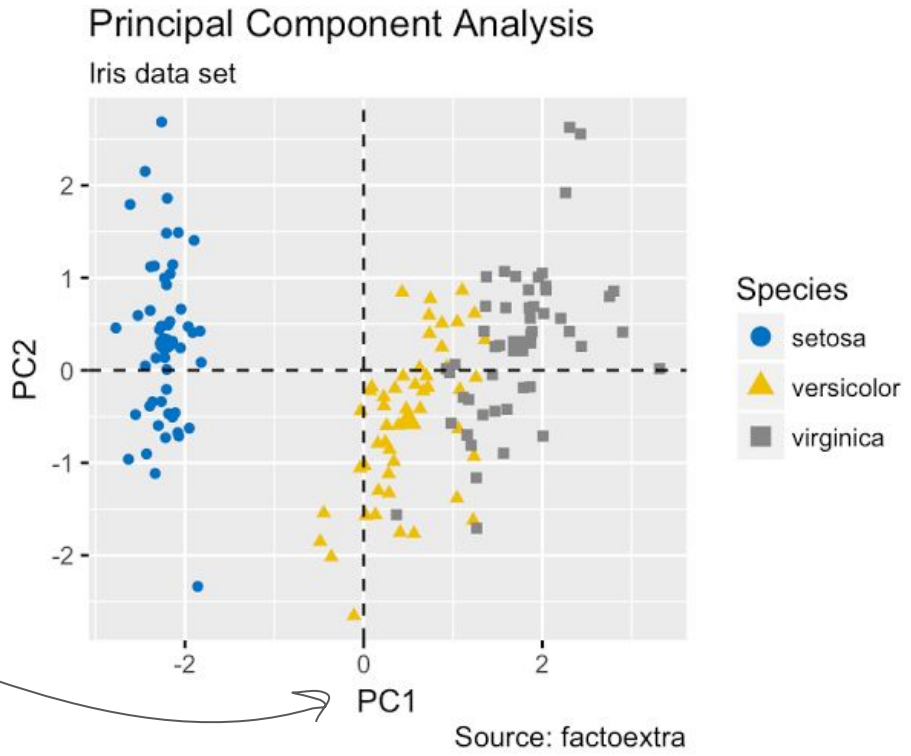
Filter

	Sepal.Length	Sepal.Width	Petal.Length	Petal.Width	Species
1	5.1	3.5	1.4	0.2	setosa
2	4.9	3.0	1.4	0.2	setosa
3	4.7	3.2	1.3	0.2	setosa
4	4.6	3.1	1.5	0.2	setosa
5	5.0	3.6	1.4	0.2	setosa
6	5.4	3.9	1.7	0.4	setosa
7	4.6	3.4	1.4	0.3	setosa
8	5.0	3.4	1.5	0.2	setosa
9	4.4	2.9	1.4	0.2	setosa
10	4.9	3.1	1.5	0.1	setosa
11	5.4	3.7	1.5	0.2	setosa
12	4.8	3.4	1.6	0.2	setosa
13	4.8	3.0	1.4	0.1	setosa
14	4.3	3.0	1.1	0.1	setosa
15	5.8	4.0	1.2	0.2	setosa
16	5.7	4.4	1.5	0.4	setosa
17	5.4	3.9	1.3	0.4	setosa

PCA

Filter

	Sepal.Length	Sepal.Width	Petal.Length	Petal.Width	Species
1	5.1	3.5	1.4	0.2	setosa
2	4.9	3.0	1.4	0.2	setosa
3	4.7	3.2	1.3	0.2	setosa
4	4.6	3.1	1.5	0.2	setosa
5	5.0	3.6	1.4	0.2	setosa
6	5.4	3.9	1.7	0.4	setosa
7	4.6	3.4	1.4	0.3	setosa
8	5.0	3.4	1.5	0.2	setosa
9	4.4	2.9	1.4	0.2	setosa
10	4.9	3.1	1.5	0.1	setosa
11	5.4	3.7	1.5	0.2	setosa
12	4.8	3.4	1.6	0.2	setosa
13	4.8	3.0	1.4	0.1	setosa
14	4.3	3.0	1.1	0.1	setosa
15	5.8	4.0	1.2	0.2	setosa
16	5.7	4.4	1.5	0.4	setosa
17	5.4	3.9	1.3	0.4	setosa



Luke Hayden
August 9th, 2018

Experiment with the code in this tutorial

Open in Workspace

MUST READ R PROGRAMMING

Principal Component Analysis in R Tutorial

In this tutorial, you'll learn how to use R PCA (Principal Component Analysis) to extract data with many variables and create visualizations to display that data.

218

EXPLORE DATACAMP'S R COURSE LIBRARY

Explore Now

Principal Component Analysis (PCA) is a useful technique for exploratory data analysis, allowing you to better visualize the variation present in a dataset with many variables. It is particularly helpful in the case of "wide" datasets, where you have many variables for each sample. In this tutorial, you'll discover PCA in R.

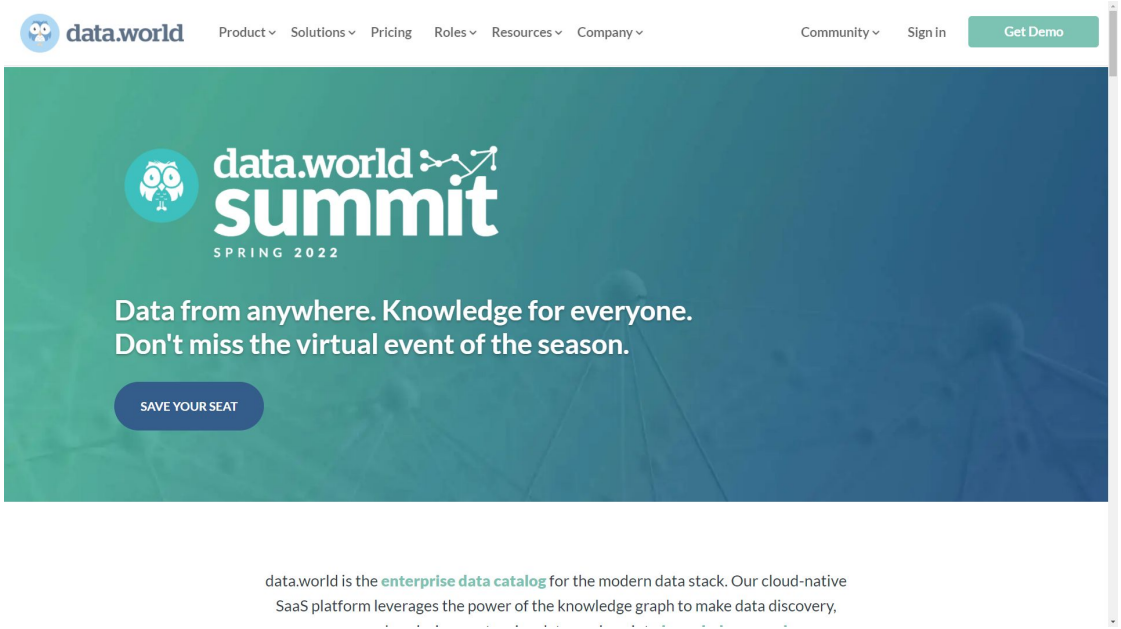
https://www.datacamp.com/community/tutorials/pca-analysis-r

```
mtcars.pca <- prcomp(mtcars[,c(1:7,10,11)], center = TRUE,scale. = TRUE)

summary(mtcars.pca)

## Importance of components:
##                           PC1    PC2     PC3     PC4     PC5     PC6
## Standard deviation     2.3782 1.4429 0.71008 0.51481 0.42797 0.35184
## Proportion of Variance 0.6284 0.2313 0.05602 0.02945 0.02035 0.01375
## Cumulative Proportion  0.6284 0.8598 0.91581 0.94525 0.96560 0.97936
##                            PC7    PC8     PC9
## Standard deviation     0.32413 0.2419 0.14896
## Proportion of Variance 0.01167 0.0065 0.00247
  ## Cumulative Proportion  0.99103 0.9975 1.00000
```

https://www.datacamp.com/community/tutorials/pca-analysis-r

Where to find **data sources**

1. data.world
2. google dataset search
3. kaggle
4. do some googling :D



Must Have Portfolios

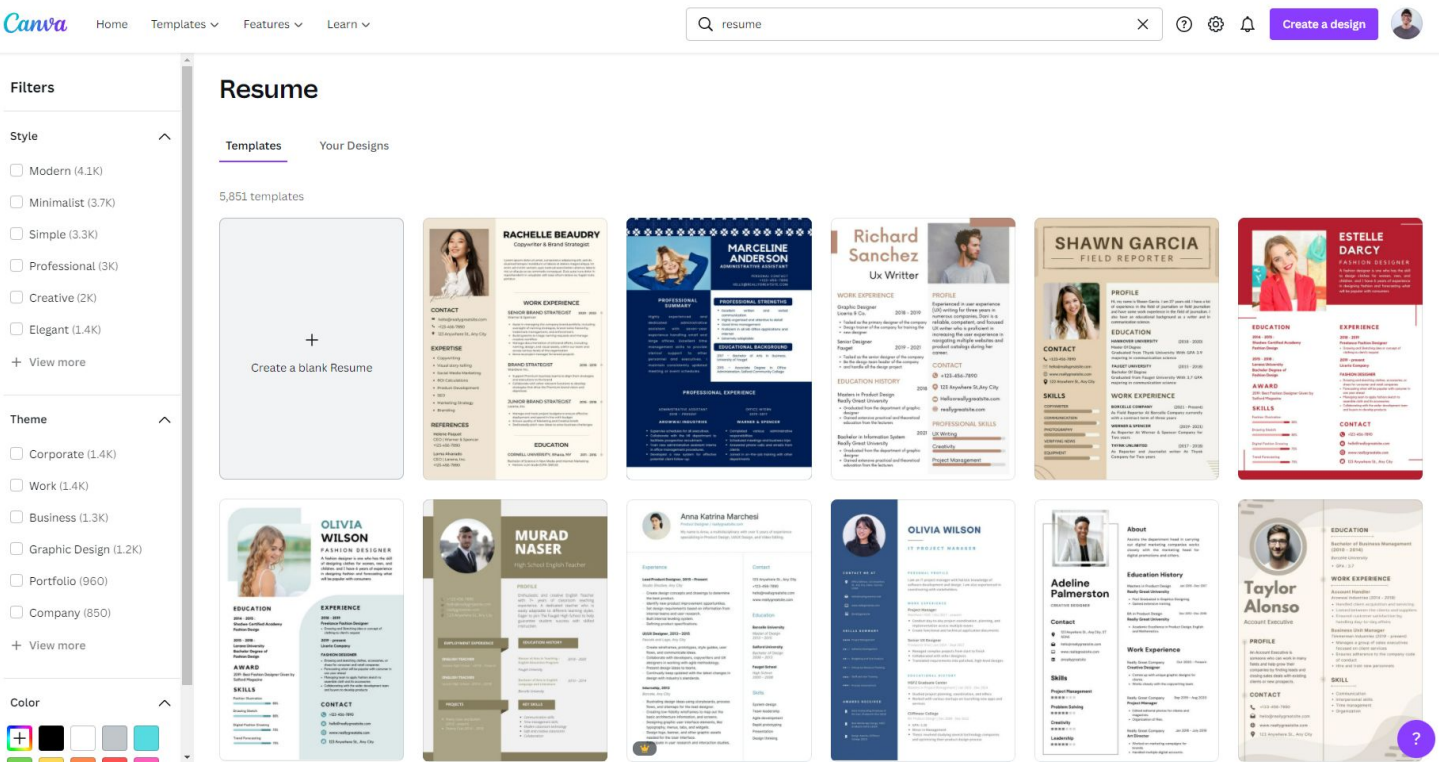
Data, Dashboard, Model, Insights

MVP (Minimum Viable Product)

- Spreadsheets
  - Simple dashboard / pivot table
- Databases SQL
  - Create database
  - Window functions/ subqueries / WITH
- Programming
  - R or Python
- Statistics
  - AB Test
  - Linear Regression
- Basic Machine Learning
  - 1x Regression
  - 1x Classification
  - 1x Clustering **or** 1x PCA
- Dashboard
  - Google Data Studio
  - Tableau **or** Power BI (using the same dataset) **or** both

Make Yourself Stand Out

- Spreadsheets
  - Simple dashboard / pivot table
- Databases SQL
  - Create database
  - Window functions/ subqueries / WITH
- Programming
  - R or Python
- Statistics
  - AB Test
  - Linear Regression
- Basic Machine Learning
  - 1x Regression
  - 1x Classification
  - 1x Clustering **or** 1x PCA
- Dashboard
  - Google Data Studio
  - Tableau **or** Power BI (using the same dataset) **or** both



https://www.canva.com/search/templates?q=resume

One page summary

Resume

1-page

CV

More than one

Your profile == first impressions
Linkedin, website, portfolio

Professional Experience
Latest one first, add more details

Education & Certifications

Skills & Portfolio

Can be switched

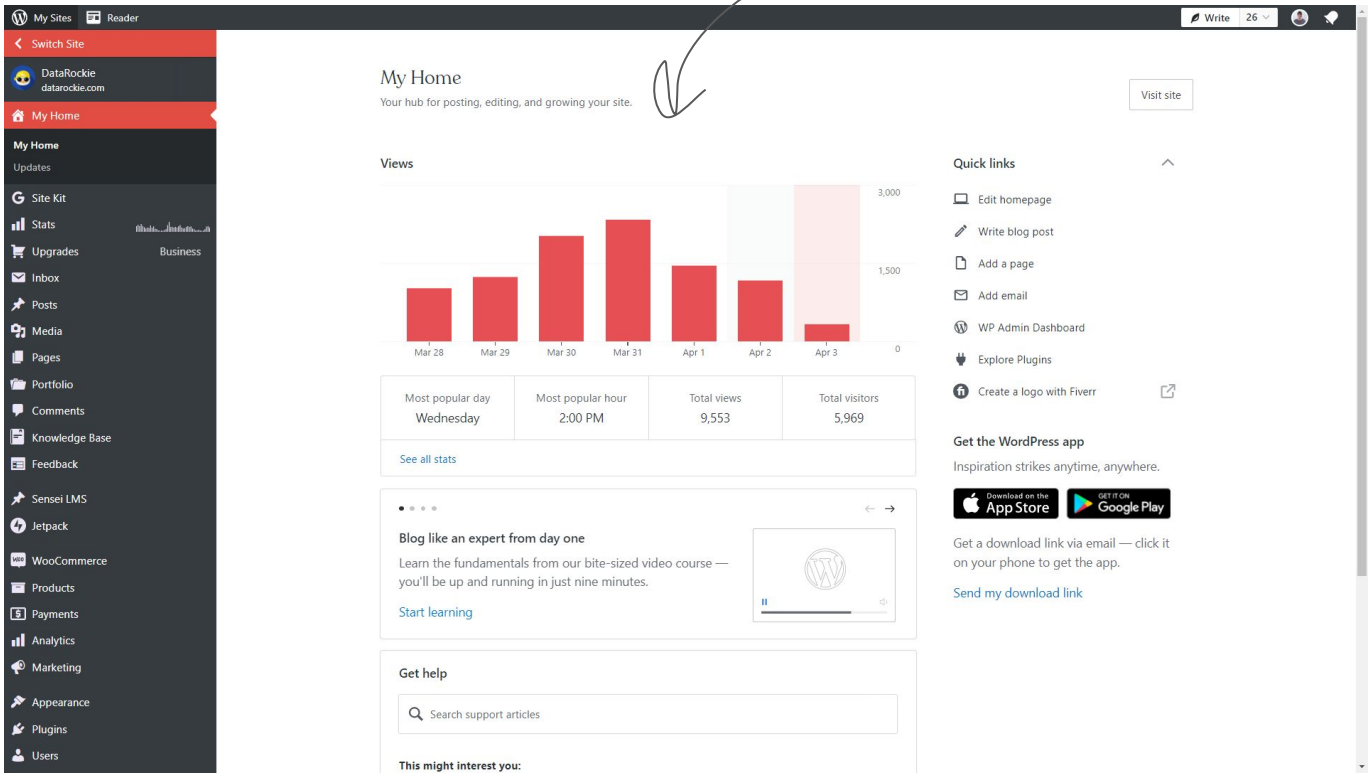
A little bit of consumer behaviour knowledge

Your resume should have the keywords used in job posts

Create Your Own Website

Data, Dashboard, Model, Insights

Full features

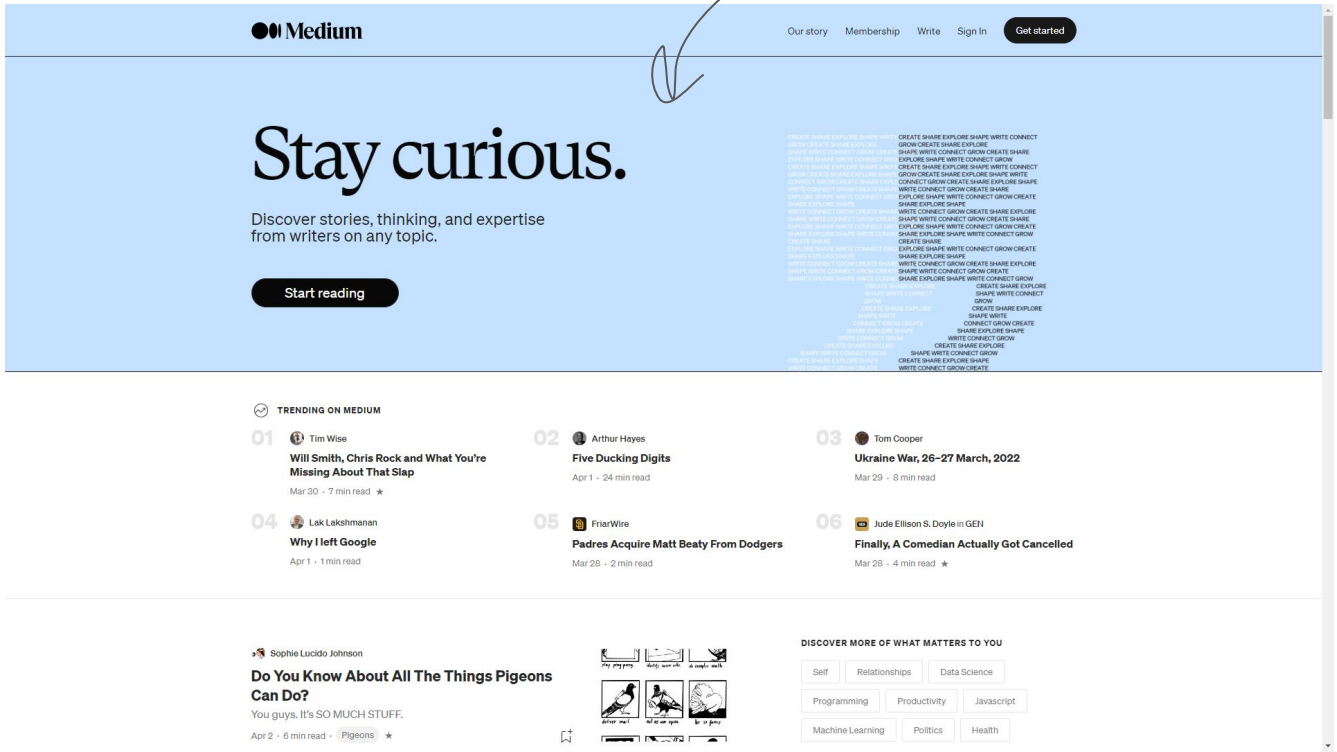


https://wordpress.com/

Create Your Own Blog

Data, Dashboard, Model, Insights

Popular blogging platform, good SEO



https://medium.com/

Build Your Online Portfolio

Data Science Bootcamp