GATE by moyyn

Data Analyst Module 1

Dr. Kavitha Chetana Didugu

BPW

Hochschule für Wirtschaft und Recht Berlin
Berlin School of Economics and Law
Startup Incubator Berlin
Das Gründungszentrum der HWR Berlin

STARTUP SCHOOL by Y Combinator

EUROPÄISCHE UNION
Europäischer Sozialfonds

be Berlin
Der Regierende Bürgermeister von Berlin
Senatskanzlei
Wissenschaft und Forschung

Things get done only if the data we gather can inspire and inform those in a position to make a difference

Mike Schmoker, School Administrator

Today’s Agenda

Your Key Takeaway

- Understanding a Data Analyst’s Day-to-day job
- Task breakdown of a Data Analysis/Analytics task
- Thinking like a Data Analyst
- Introduction to the first real-life Project

What exactly is Data Analytics?

Definitely not just some lines of Code!!

What does a Data Analyst actually Do?

Again, not just some lines of Code!!

Understand the Problem

What is the Objective? What decisions or outcomes is expected from the data?

Understand the Data

What does the data look like? Is there any relattion within the various data fields/variables?

Visualise the Right Data

What is the best way to represent the data so that the audience can follow the story

Data Story Telling and Dashboarding

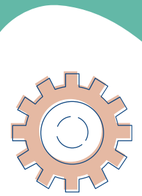
After going through heaps of data, you need to make a meaningful narrative out of it for your stakeholders

What should a GOOD Data Analyst know?

Hard Skills

Probability and Statistics

Descriptive Statistics, Predictive Statistics, Hypothesis Testing



Data Interpretation, Logic and Reasoning

Understanding how to read and understand inputs, outputs, Objective functions, etc.

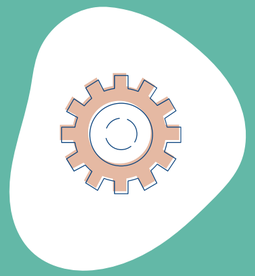
Coding

Basic skills in R, Python or Julia

Soft Skills

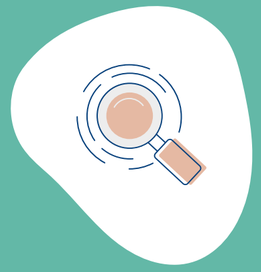
Communication skills

Data Analyst work in cross-functional teams, and need to communicate complex mathematical relationships in simple English



Data storytelling

From large datasets, only the useful insights need to communicated. A Data Analyst should figure out what is useful and



Dashboarding

Python (Plotly and Dash, Seaborn, Matplotlib), Excel, Tableau, Power BI

The Path to being a Great Data Analyst

This training is just the beginning!

Read Up!

There are lots of articles and blogs just a Google search away

Practice!!

Kaggle.com has numerous datasets which you can analyse, visualise and dashbboard. Practice, code, learn from other participants

Be Curious!!!

Watch a lot of Youtube videos, if you get stuck, read up exchange forums such as stackoverflow

Interaction

Anatomy of a Data Analytics Project

Step-by-Step Task Breakdown

Anatomy of a Data Analytics Project

- Problem Comprehension or Definition
- Exploratory Data Analysis
- Data Visualisation
- Data Storytelling: Finding and Presenting Insights
- Reporting and Dashboarding

Real life example

Customer engagement data for 2020 and 2021 for a Department store

Given the data about customer engagement for years 2020 and 2021, your job as a Data Analyst is to provide insights to help the Department store figure out the conversion rate (Conversion rate is the proportion of views that actually turned out to be purchases). Additionally, figure out if customer conversion rate has improved or degraded, and the reason why.

Link to data: https://www.kaggle.com/datasets/mkechinov/ecommerce-events-history-in-electronics-store

Think like a Data Analyst

1. What am I looking for? Define the Problem
2. How is each variable in the dataset related?
3. What is the best way to visualise this relationship?
4. What does this relationship mean in the business context?
5. What are the recommendations I can give based on the above observations?
6. How do I best present these recommendations?

Let's look at some code: https://tinyurl.com/kaggleda

Further Reading and Resources

Learn more about:
Pandas: https://www.kaggle.com/learn/pandas
Data Cleaning: https://www.kaggle.com/learn/data-cleaning
Data Analysis and Statistics: https://www.khanacademy.org/kmap/measurement-and-data-g/md220-data-and-statistics
https://app.datacamp.com/learn/courses/exploring-and-analyzing-data-in-python
https://assessment-v2.datacamp.com/analytic-fundamentals-theory
Creating Visualisations: https://seaborn.pydata.org/tutorial.html
https://matplotlib.org/
https://pandas.pydata.org/docs/user_guide/visualization.html
Dashoboarding: https://plotly.com/python/

For further inspiration:
https://plotly.com/examples/
https://streamlit.io/gallery?category=data-visualization

GATE

by moyyn

German Academy for Technology and Entrpreneurship

BPW

Hochschule für Wirtschaft und Recht Berlin
Berlin School of Economics and Law

Startup Incubator Berlin
Das Gründungszentrum der HWR Berlin

STARTUP SCHOOL
by Y Combinator

EUROPÄISCHE UNION
Europäischer Sozialfonds

be Berlin

Der Regierende Bürgermeister von Berlin
Senatskanzlei
Wissenschaft und Forschung

Feedback Survey

http://www.moyyn.com/gate-feedback