

DATA PROJECT



#6



Twitter Data



FABIAN WERKMEISTER

DATA

WHAT data

&

WHERE to get it



DATA SOURCE

Tweets on climate change

- **Text body of the tweet**
- **Sentiment of the tweet**
 - 0 → neutral
 - 1 → beliefs in climate change
 - -1 → no belief in climate change
 - 2 → factual news



You can find & download
the dataset on Kaggle



LINK IN NOTION



FABIAN WERKMEISTER

DATA TRACKS

SELECT ONE or COMBINE



I - Data Engineering

II - Data Analytics

III - Data Science



TRACK I

DATA ENGINEERING

Clean & Preprocess tweets

- Remove stop words, special characters, emojis
- Apply text normalization (stemming, lemmatization)
- Extract hashtags per tweet



TUTORIALS IN NOTION



FABIAN WERKMEISTER

TRACK I DATA ENGINEERING



Search for a trending
climate hashtag on Twitter

Use it to get additional
tweets from the Twitter API



TUTORIALS IN NOTION



FABIAN WERKMEISTER

TRACK II

DATA ANALYTICS

Text Analytics

- Calculate descriptive statistics for hashtags
- Calculate tweet lengths and compare these lengths by sentiment
- Identify links in tweets and calculate the number of links per tweet



TRACK II

DATA ANALYTICS



Do tweets of a particular sentiment show specific characteristics?

Check for correlations between sentiment and text length & no. of links



FABIAN WERKMEISTER

TRACK III

DATA SCIENCE

Sentiment Prediction (NLP)

- Prepare the tweets to be fed into a model
 - BoW
 - TF-IDF
 - Embeddings
- Try different models to predict the sentiment of a tweet



TUTORIALS IN NOTION



FABIAN WERKMEISTER

TRACK III

DATA SCIENCE



Use NER to extract
locations, persons and
organizations

Are certain persons often
associated with
a specific sentiment?



FABIAN WERKMEISTER

LEARNINGS



why YOU



should do



this project



FABIAN WERKMEISTER

LEARNINGS

Track I - Data Engineering
You learn how to work with & preprocess text data

 **Track II - Data Analytics** 
You learn how to analyze text data & derive statistics from it

Track III - Data Science
You learn how to prepare text data for machine learning



**WHAT WOULD
YOU DO WITH
THIS DATA?**

**SEE YOU IN
THE COMMENTS**



FABIAN WERKMEISTER

HI, I'M FABIAN



**I HELP YOU
TO BECOME
FLUENT IN DATA**



FOR PROJECT IDEAS EACH FRIDAY

+ Follow

& hit the 



FABIAN WERKMEISTER

 **BONUS** 



FREE Data



Notion Space



FABIAN WERKMEISTER

DATA NOTION SPACE

Fabian Werkmeister

I help you to become fluent in Data and enable you to translate between business & tech ★ Data Scientist @FUNKE ★  Python
 Data Science  Data Engineering

Talks about #data, #python, #mindset, #datascience, and #productivity

Hamburg, Hamburg, Germany · [Contact info](#)

[FREE Data Notion Space](#) 



✓ All Project Resources

✓ Project Management Tools

✓ Tutorials & Learning Roadmap



FABIAN WERKMEISTER

DATA NOTION SPACE

Simply click "Duplicate" and get
your very own Notion Data space

The screenshot shows a Notion workspace interface. At the top, there is a search bar with the word 'Search' and a 'Duplicate' button, both circled in orange. To the right of the search bar is the Notion logo. A large orange arrow points downwards from the search bar area towards a grid of nine cards. The cards represent different steps in a data science process:

- 1 Project Structure:** An image of a building under construction with cranes.
- 2 Data Collection:** An image of hands holding chestnuts next to a bucket.
- 3 Data Cleaning:** An image of a hand holding a yellow spray bottle.
- 4 Data Exploration & Analysis:** An image of people looking at a map.
- 5 Preprocessing & Feature Engineering:** An image of various tools like pliers and a hammer.
- 6 Modeling & Evaluation:** An image of several small robots.
- 7 Productionize Model:** An image of a large cargo ship.
- 8 Showcase Results:** An image of a clapperboard.
- 9 Summary & Retrospective:** An image of a sunset sky.

