

# DATA PROJECT



#18



EMPLOYEE  
TURNOVER



FABIAN WERKMEISTER

# **DATA**

---

**WHAT** data

&

**WHERE** to get it



# DATA SOURCE

Fictional data about  
employee terminations

- **age**
- **gender**
- **city**
- **business unit**
- **termination date & reason**



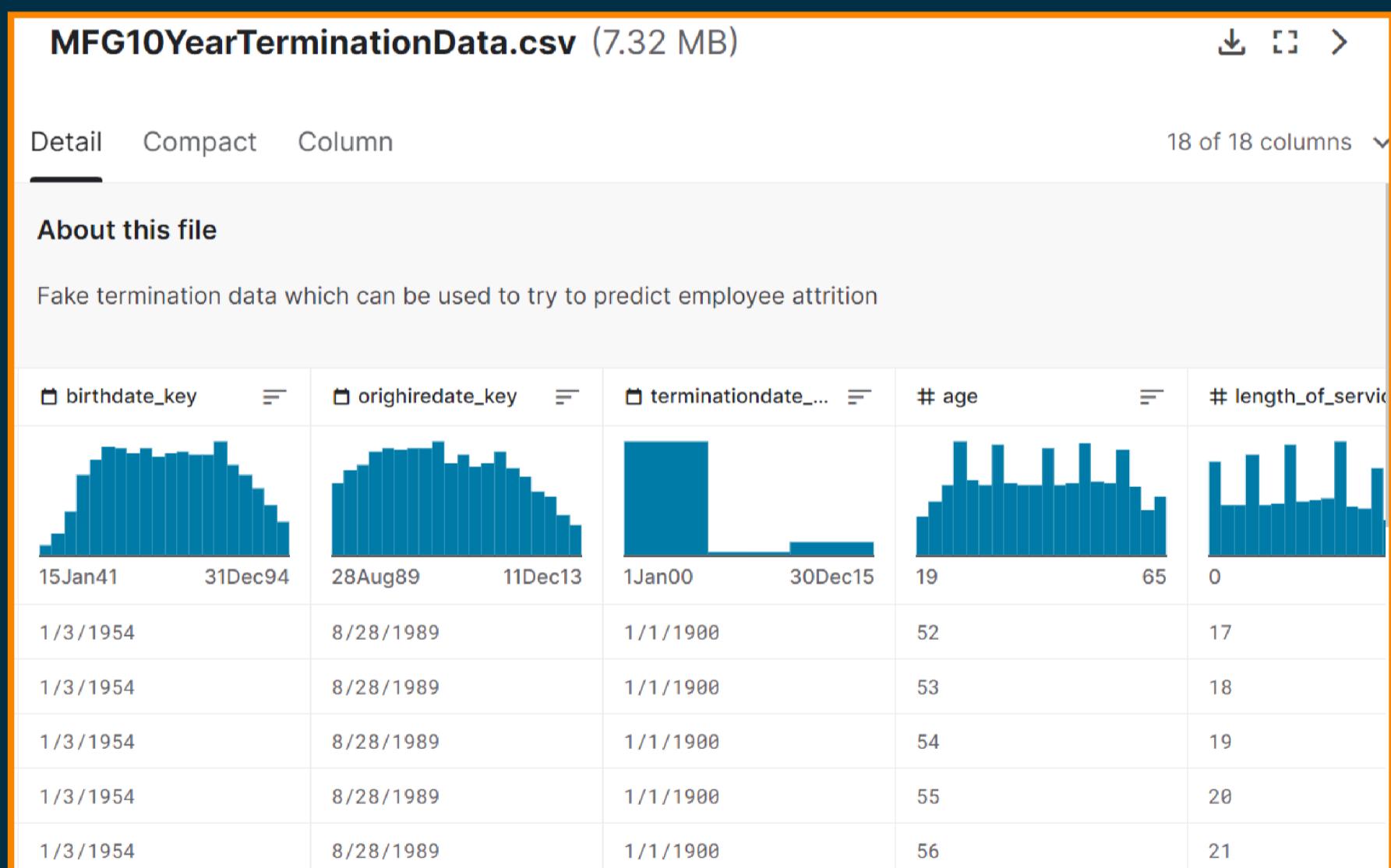
BE CAUTIOUS --- THERE ARE MULTIPLE ROWS FOR  
ONE EMPLOYEE, WITH EACH ROW REPRESENTING  
ONE CONSECUTIVE YEAR OF DATA COLLECTION!



FABIAN WERKMEISTER

# DATA SOURCE

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

Transform data into a  
panel data format

- **Each row should represent one employee**
- **Convert yearly changing values into a yearly column**  
(e.g. "age\_2012", "age\_2013", etc.)
- **Keep also other columns**  
(e.g. birthdate will always be the same)



TUTORIALS IN NOTION



FABIAN WERKMEISTER

# TRACK I DATA ENGINEERING



The fictional data only  
ranges from 2006 to 2015

Create additional fictional  
data up until May 2023



TUTORIALS IN NOTION



FABIAN WERKMEISTER

# TRACK II

# DATA ANALYTICS

## Employee Analytics

- Demographical distribution over the years
- Job & department distributions over the years
- Hiring and terminations over the years



# TRACK II

# DATA ANALYTICS



Analyze employee  
turnover causes over time

Do termination reasons  
correlate with e.g. age,  
department, job title or  
length of service?



TUTORIALS IN NOTION



FABIAN WERKMEISTER

# TRACK III

# DATA SCIENCE

## Cluster Employees

- Use age, length of service, department & job title for the year 2015
- Repeat this for other years
- Is the same person in the same cluster each year or do they "switch"?



TUTORIALS IN NOTION



FABIAN WERKMEISTER

# TRACK III

# DATA SCIENCE



Build a churn model, which  
predicts the total employee  
turnover for a given year

Adapt this and try to  
predict, which person will  
probably quit their job



TUTORIALS IN NOTION



FABIAN WERKMEISTER

# **LEARNINGS**

---



**why YOU**



**should do**



**this project**



FABIAN WERKMEISTER

# LEARNINGS

**Track I - Data Engineering**  
**Data Modeling & Transformation,**  
**Creating Fictional Data**



**Track II - Data Analytics**  
**Descriptive Analytics for**  
**Longitudinal Data, Correlation**  
**Analysis**



**Track III - Data Science**  
**Cluster Analysis, Churn Prediction**



**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 placeholder 'Search' and a 'Duplicate' button, both of which are circled in orange. An orange arrow points downwards from this bar towards a grid of nine cards. The cards are arranged in three rows of three. Each card contains an icon and a numbered title:

- 1 Project Structure
- 2 Data Collection
- 3 Data Cleaning
- 4 Data Exploration & Analysis
- 5 Preprocessing & Feature Engineering
- 6 Modeling & Evaluation
- 7 Productionize Model
- 8 Showcase Results
- 9 Summary & Retrospective



FABIAN WERKMEISTER