

Group 23: Twitter Sentiment Analysis

Key Ideas:

- Machine Learning on Big Data
- **10** sentiments
- **31,130,257** (english) Tweets
- **3.01 GB**

Key Tasks:

- Preprocessing of Tweets
- Hashing TF and IDF
- Training a Deep Neural Network Model for Classification
- K-fold Cross Validation
- Grid Search/Hyperparameter Tuning

Tools and Technologies Used:

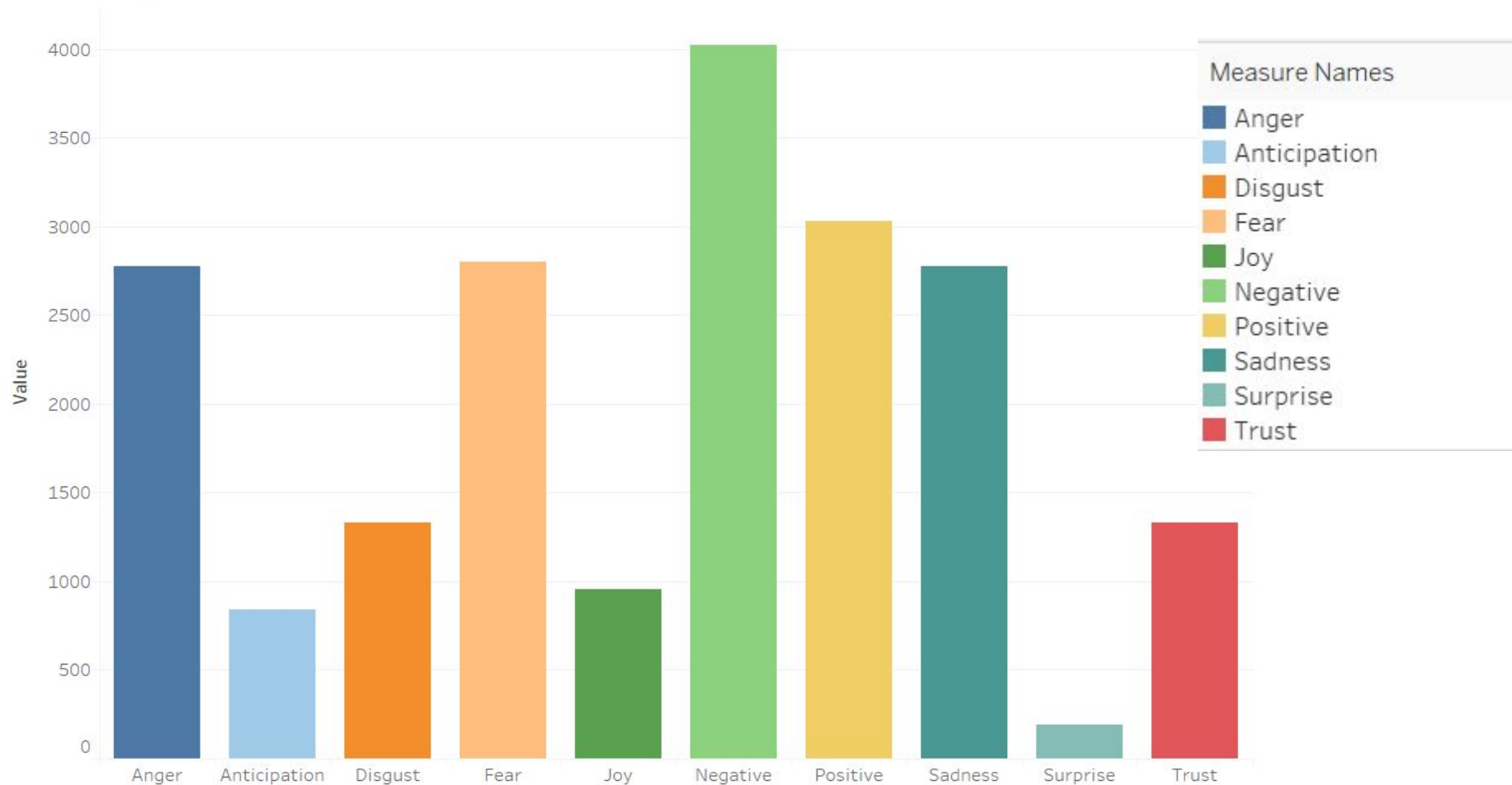
- Azure Databricks
- Azure SQL Database Warehouse
- Apache Spark with Scala
- Spark MLLib (Machine Learning)
- Spark SQL (Data Management)
- Tableau (Data Visualization)



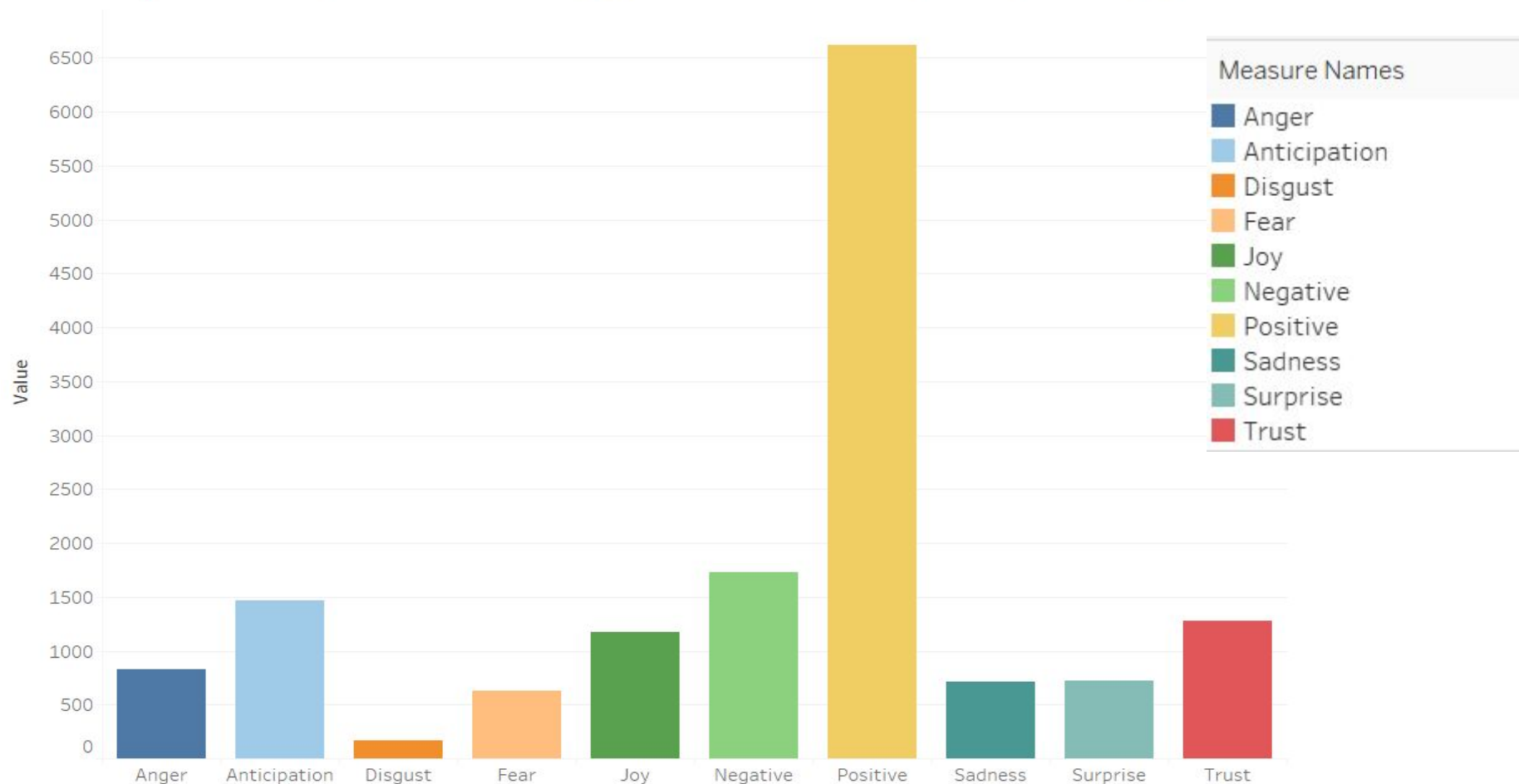
Azure Databricks



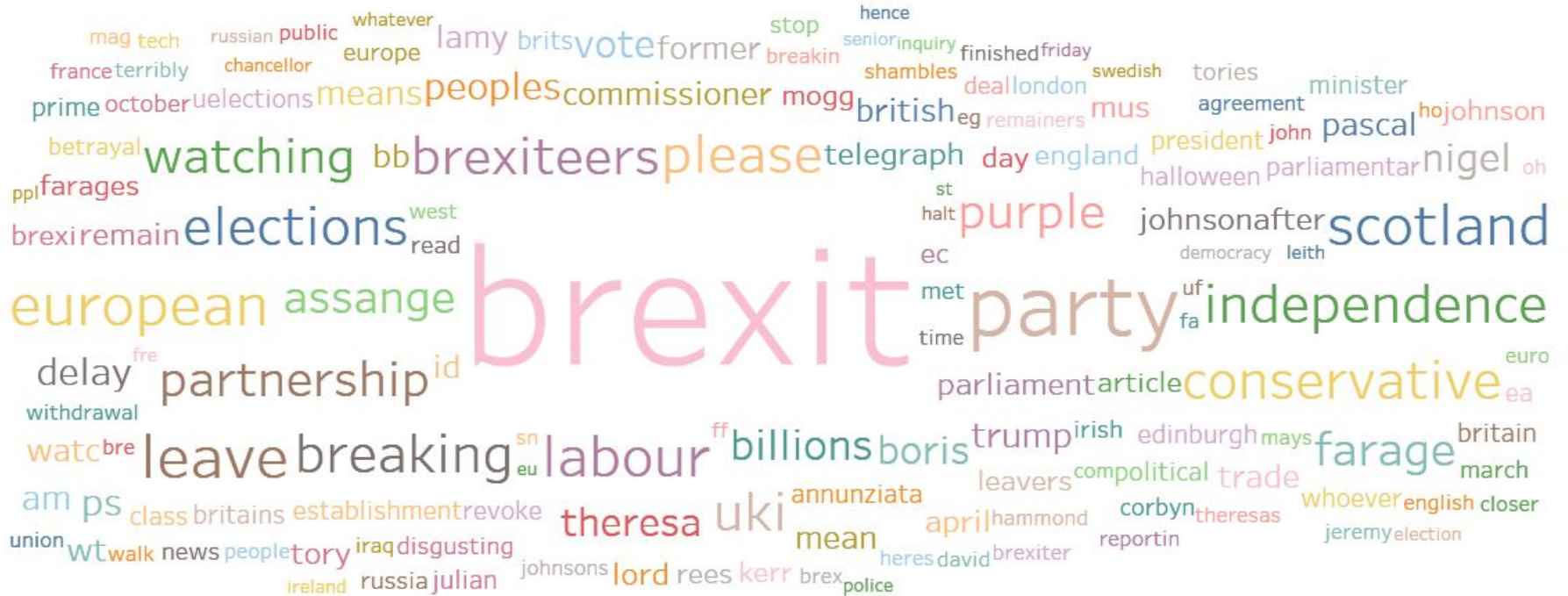
Example 1: Number of **#brexit** Tweets for each Sentiment



Example 2: Number of **#got** Tweets for each Sentiment



Example 1: Word Cloud for **#brexit** Tweets



Example 2: Word Cloud for #got Tweets

