Salary Prediction

Regression Analysis

Overview

Data Collection

Data Exploration - ETL

Regression Model

Conclusion

Data Collection

Source:

- <u>Kaggle</u>
- Scraped from Glassdoor.com
- 2017 dataset



Data Exploration

- Extract, Transform and Load ETL
- Data Visualizaton- Tableau
- Descriptive Statistics

Finding
Duplicates

Focusing on Columns Needed

Relevant columns for the analysis

Splitting and
Deleting special
characters

Cleaning up each column with special values or characters

Exploring each columns

Unique values of each column



Questions

- What's the median pay for roles in the field of data??
- What's the top-paying job title in the data-driven realm?
- Which technical skills are the hottest in the job market?
- Which state has the highest paying data related jobs?
- Is my paycheck influenced by the company's size and industry sector?
- Does higher salary mean higher job satisfaction?

Conclusion:

97K

Median Salary

Python

Hottest technical skill

California

Highest paying data related jobs

CHALLENGES

- Data Collection: Choosing the right data
- Model: Choosing the right model

Thank you!

This is just the beginning!