

# Web Scraping Indeed Data Science Positions

---

Aakash Tandel  
General Assembly - DSI 5

# Mission:

Determine how title,  
skills, and locations  
affect data science  
salaries

# Overview

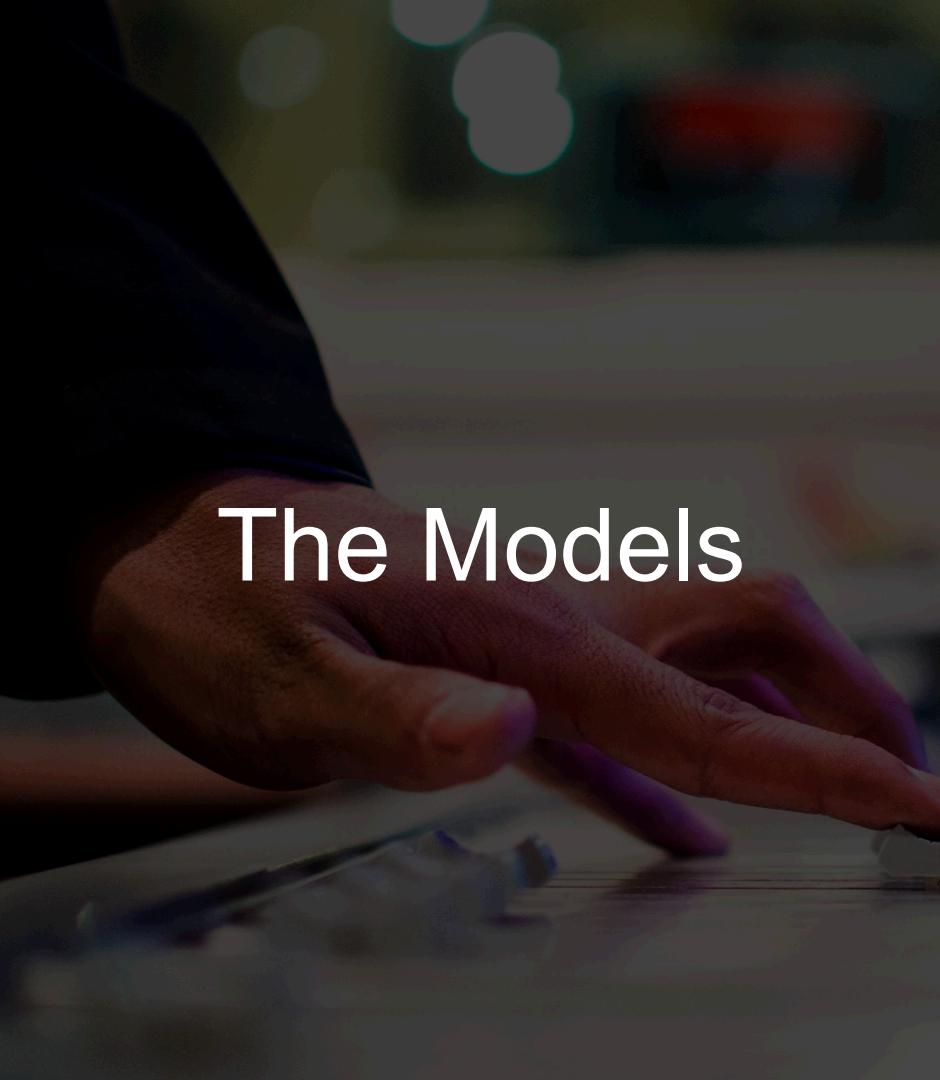
---

- The Data
- The Models
- The Results
- The Recommendations

# The Data

---

- Web Scrapped Job Postings from Indeed.com
- 180,000 results pulled
- Information from over 70 cities
- Median Salary: \$86,711



# The Models

Classification Models - above or below the median salary

Random Forest, XGBoost,  
Logistic Regression, and  
Support Vector Machines

Accuracy of around 70%

---

# The Results

---

Above Median - Data Scientist, Machine Learning

- Below Median - Data Analyst, Research Analyst
- Above Median Locations - MA, IL, WA, and DC
- Below Median Locations - AZ, NM, and IN

# The Recommendations:

- Scope of Work
  - Data Analyst versus Data Scientist
- DC Data Science Salaries

# Summary



# Questions

