THE MOST AFFORDABLE DESTINATION FOR DATA SCIENTISTS

XIAO HUANG

1/10/20

TOOLS

Beautifuloup





Postgre**SQL**



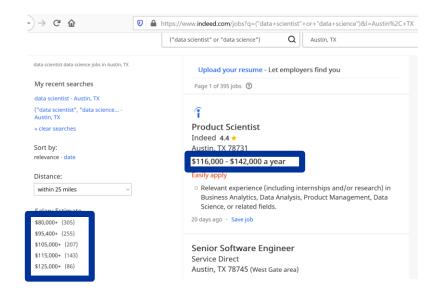


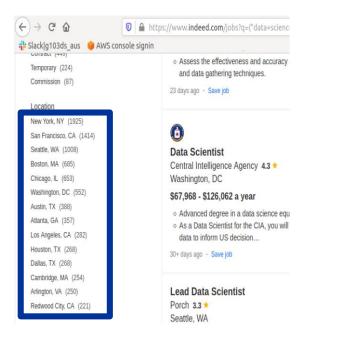






SciPy

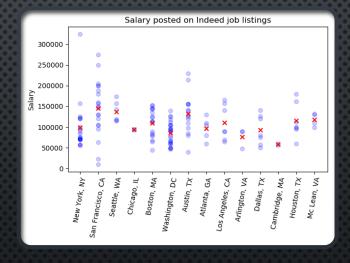


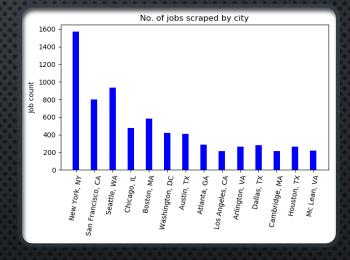


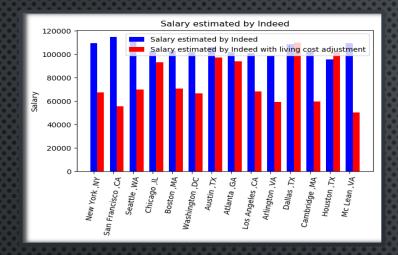
DATA SOURCE: WEB SCRAPED ON INDEED

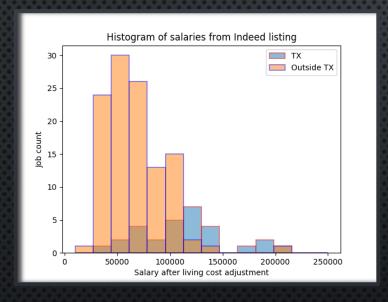
CHALLENGE: SPARSE DATA

- TOTAL NUMBER OF JOBS SCRAPED: 14000+
- Total number of Jobs Have salary posted: ~300
- 9/14 MAJOR CITIES HAVE LESS THAN 10 SAMPLES









EDA AND HYPOTHESIS TESTING

- EDA: CITIES IN TEXAS SIMILAR SALARY AND LIVING COST
- Data grouped as in TEXAS and in other states
- A/B TESTING: SALARY AFTER LIVING COST ADJUSTMENT IN TEXAS
 IS 34K HIGHER THAN NOT IN TEXAS WITH P VALUE 0.04

CONCLUSION AND FUTURE STUDY

• Salary adjusted by Living Cost in Texas are \$34k higher than in other states

- FUTURE STUDY:
 - Consider other factors like weather, crime, amenities etc.
 - ADJUST WEIGHTS OF FACTORS

CONTACT INFO

- LINKEDIN: HTTPS://WWW.LINKEDIN.COM/IN/XIAO-HUANG-4A545B87
- GITHUB REPO: HTTPS://GITHUB.COM/IVBOH/MOST_AFFORDABLE_DESTINATIONS_DATA_SCIENTIST



Ph.D. in Statistics, Washington University in St. Louis B.S. in Statistics, University of Science and Technology of China