

The Most Affordable Destinations for Data Scientists

Xiao Huang

1/10/20

Overview

TOOLS

EC2 Web scraper

BeautifulSoup parser

Postgres SQL

Python matplotlib stats

DATA

indeed.com

areavibes.com

Data source: Popular Cities on Indeed

Slack|g103ds_aus AWS console signin

Contract (443)

Temporary (224)

Commission (87)

Location

- New York, NY (1925)
- San Francisco, CA (1414)
- Seattle, WA (1008)
- Boston, MA (685)
- Chicago, IL (653)
- Washington, DC (552)
- Austin, TX (388)
- Atlanta, GA (357)
- Los Angeles, CA (282)
- Houston, TX (268)
- Dallas, TX (268)
- Cambridge, MA (254)
- Arlington, VA (250)
- Redwood City, CA (221)

Assess the effectiveness and accuracy and data gathering techniques.

23 days ago · [Save job](#)

Data Scientist

Central Intelligence Agency 4.3 ★

Washington, DC

\$67,968 - \$126,062 a year

- Advanced degree in a data science equ
- As a Data Scientist for the CIA, you will data to inform US decision...

30+ days ago · [Save job](#)

Lead Data Scientist

Porch 3.3 ★

Seattle, WA

Company reviews Find salaries

what
job title, keywords, or company

where
city, state, or zip code

Austin, TX

Data Scientist

Shop LC ★★★★★ 14 reviews - Austin, TX 78728

Apply Now

Hi there! Thanks for considering our Data Scientist role.

As our Data Scientist, you will be responsible for all advanced statistical methods utilized to solve a variety of business problems in support of sales, marketing, affiliate relations, and customer service activities across all businesses. You must be able to clearly communicate with business leaders, identify and understand business problems, create solutions using analytic models and use the models to help make data driven decisions. You must be able to extract insights from the analytics models as well as create and articulate business stories from this data.

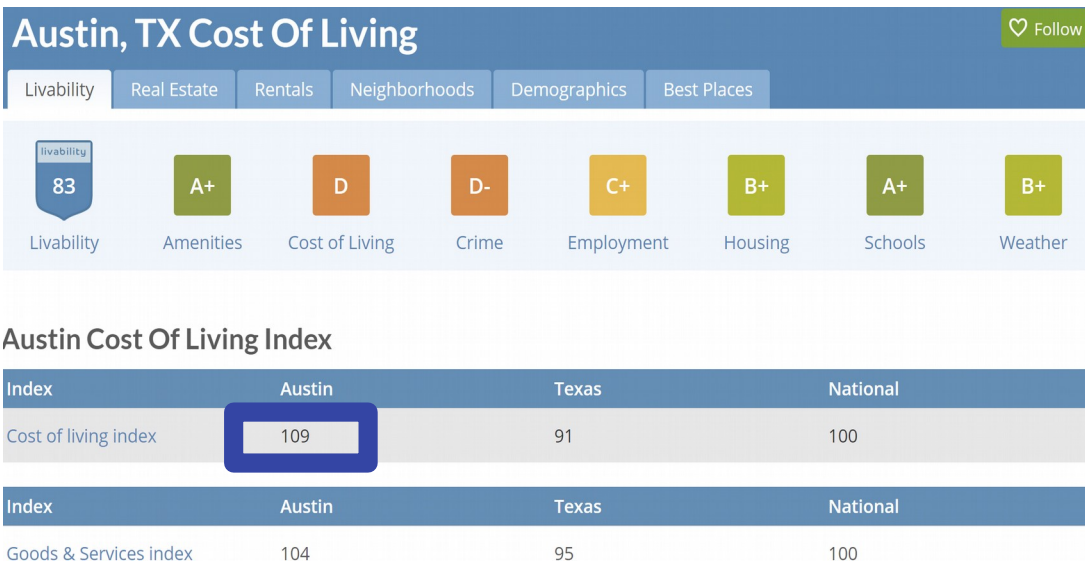
What you'll be doing in this role:

- Building, maintaining and analyzing customer data
- Applying statistical methods to identify and execute means of optimizing customer

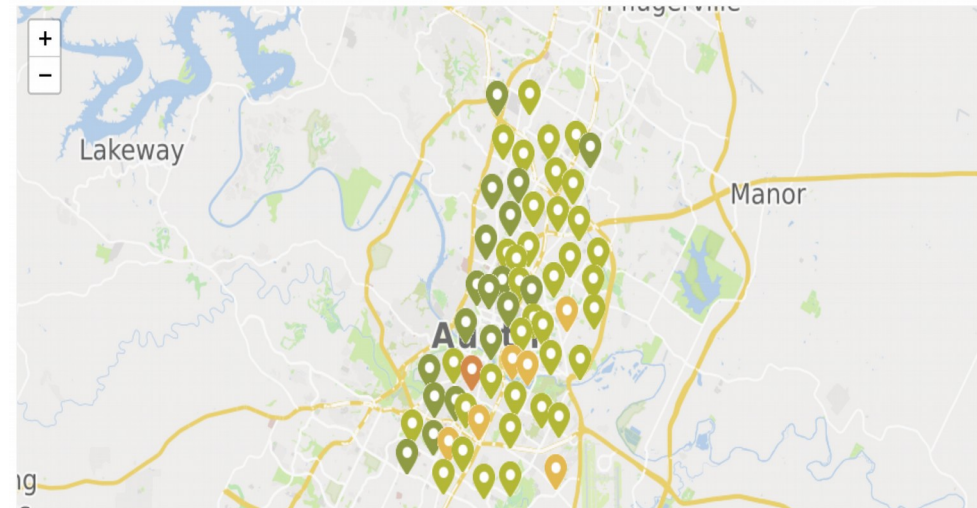
Data source: Living Cost on Areavibes

What's considered: groceries, housing, utilities, transportation, goods & services and health care

What's not considered: housing price & tax,

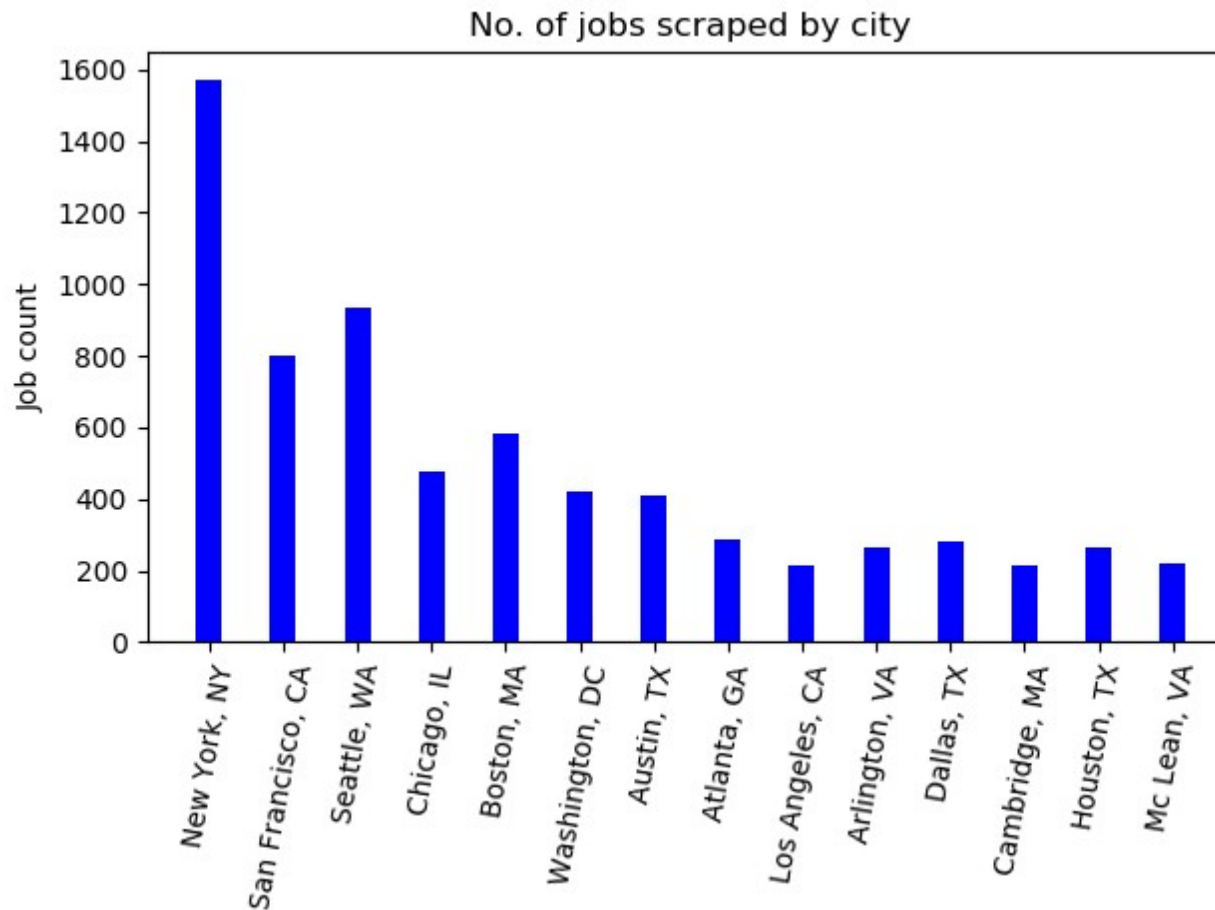


Austin Neighborhoods Map

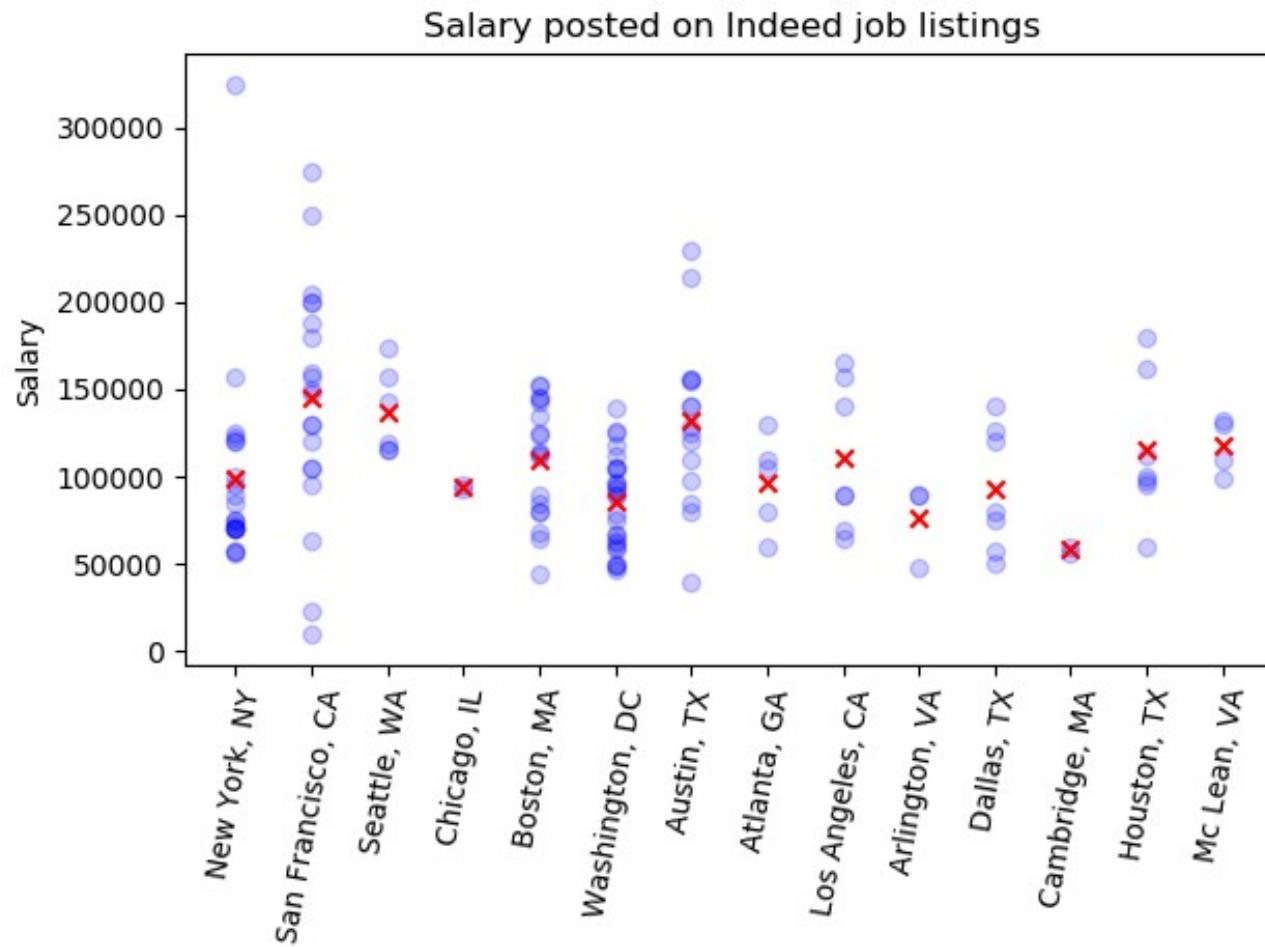


Source: The Austin, TX cost of living data displayed above is derived from the Council for Community and Economic Research (C2ER) for 2019.

Data source: Indeed Job Listing



Challenge: Sparse Data



Additional source: Indeed

The screenshot shows the Indeed job search interface. At the top, the browser address bar displays the URL: `https://www.indeed.com/jobs?q=("data+scientist"+or+"data+science")&l=Austin%2C+TX`. Below the address bar, the search query is `("data scientist" or "data science")` and the location is `Austin, TX`.

On the left side, under the heading "data scientist data science jobs in Austin, TX", there are sections for "My recent searches" (listing "data scientist - Austin, TX" and a more specific search with a "clear searches" link), "Sort by:" (set to "relevance - date"), and "Distance:" (set to "within 25 miles").

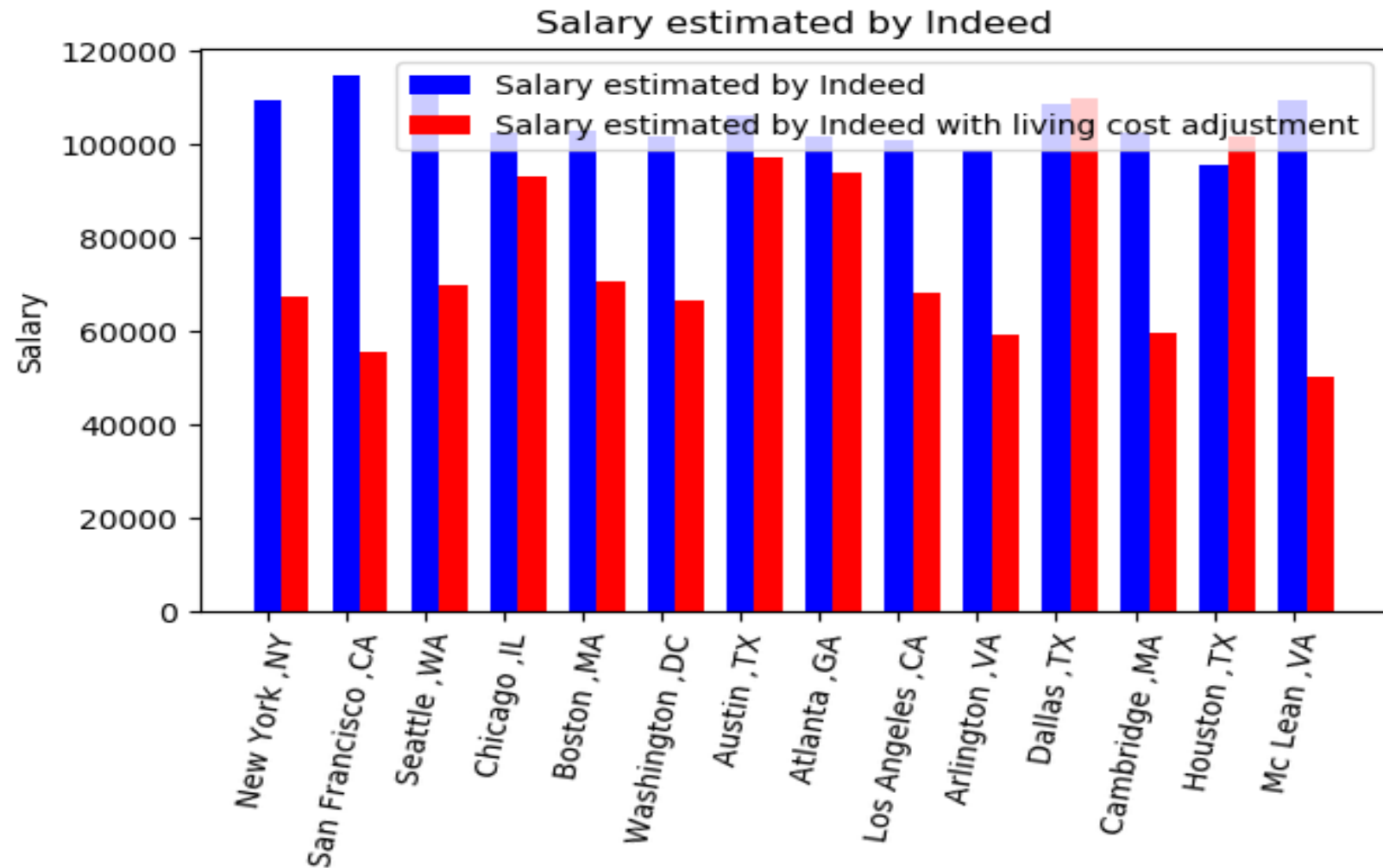
A sidebar titled "Salary Estimate" is highlighted with a blue border. It lists salary ranges and the number of jobs for each:

- \$80,000+ (305)
- \$95,400+ (255)
- \$105,000+ (207)
- \$115,000+ (143)
- \$125,000+ (86)

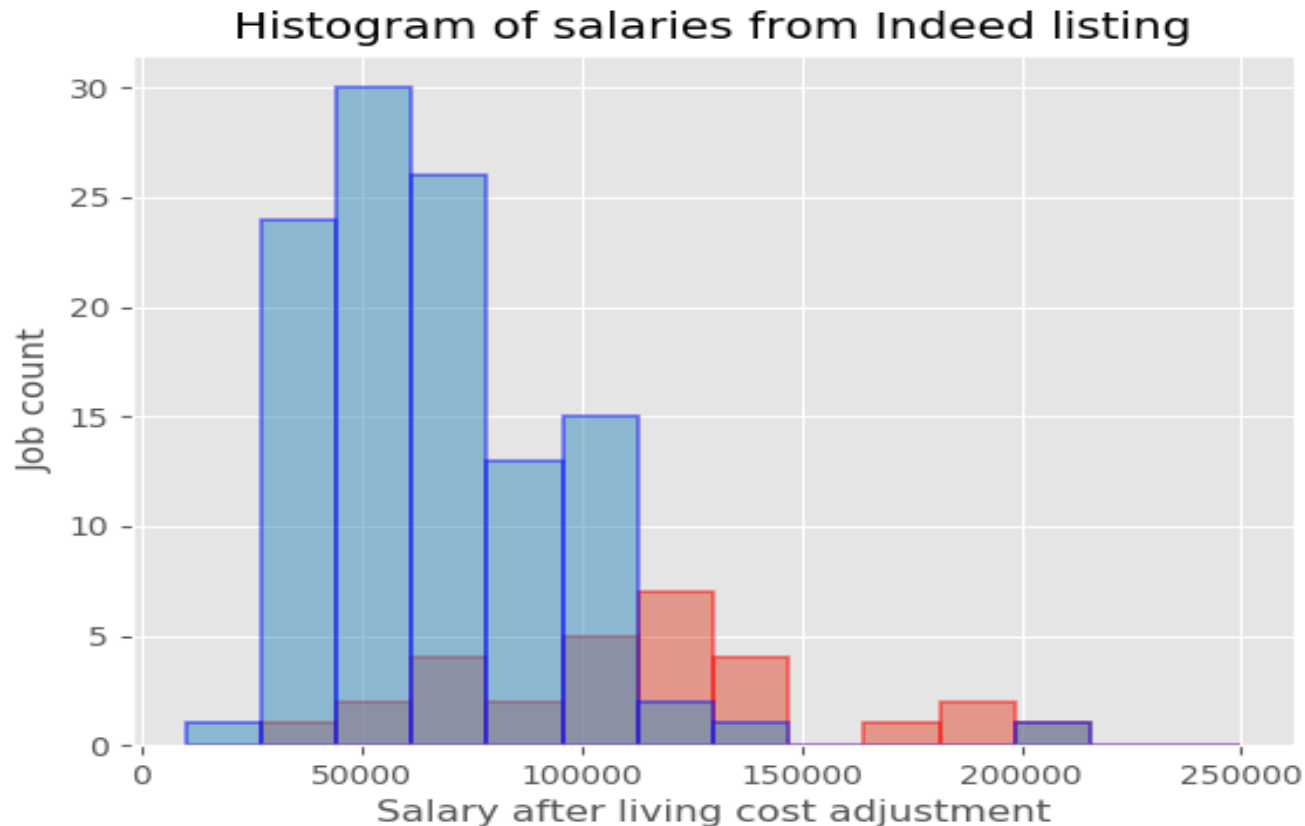
The main job listings area shows "Page 1 of 395 jobs". The first listing is for a "Product Scientist" at Indeed, with a 4.4-star rating, located in Austin, TX 78731, with a salary range of \$116,000 - \$142,000 a year. It includes an "Easily apply" button and a list of requirements: "Relevant experience (including internships and/or research) in Business Analytics, Data Analysis, Product Management, Data Science, or related fields." The listing was posted 20 days ago and has a "Save job" link.

The second listing is for a "Senior Software Engineer" at Service Direct, located in Austin, TX 78745 (West Gate area).

EDA: Combining Salary and Living Cost



Hypothesis testing: TX vs Outside



H0: TX = Non_TX Ha: TX > Non_TX one tail t test p value = 1e-06

H0: TX - Non_TX = 34k Ha: TX - Non_TX > 34K One tail t test p value = 0.041

95% confidence interval for difference of mean = (30950, 65105)

Conclusion and future study

Conclusion:

Popular destinations for data scientists in Texas are more affordable than those outside Texas. (Ask for 34k more for the similar position)

Future study:

- 1) Consider other factors like weather, crime, amenities and etc.**
- 2) Analyzing the detailed description to find correlation within and between qualifications and the types of industries.**