Salary Dataset Analysis:
An in-depth exploration

By: Reese , Kshitij, Vinh, & Swikar

Background, Objective, and Goals

The dataset in this project is based on data-related jobs in the US. The following characteristic will be the focus of this analysis

Key Performance Indicator: *Paid Wage Per Year*

The objective for this analysis is to breakdown roles by desirable characteristics in the industry that a worker may consider based on the following:

Employer Name   *Education Level*

College Major   *Experience*

Prevailing Wage   *Work City/State*

Job Title Subgroup.

The goal is to explore these characteristics using data analysis to provide answers for entry-level or existing industry professionals through asking high-level question. This is accomplished by Preparing, Processing, Analyzing, and Visualizing data. This creates a more transparent data story for stakeholders and interested parties.

Question #1: Do specific sub-types of data-related jobs have higher or lower salaries than others?
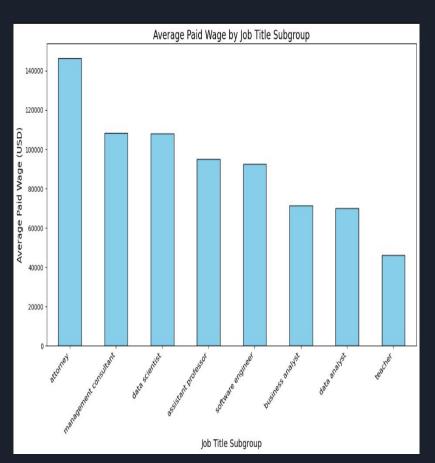
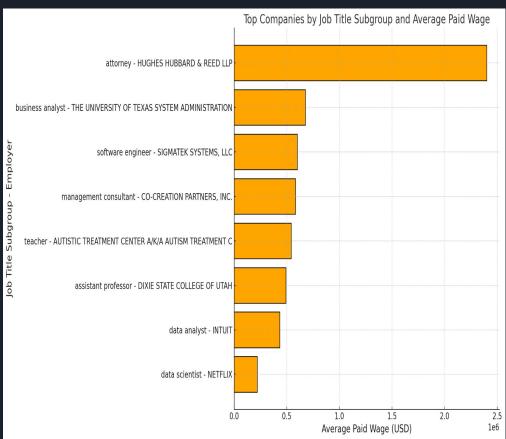1a - Which companies have the highest salaries for those subtypes ?

1b - Changes with location of the job ?

1c - Will the answer change if I take the standard of living into account ?

Q1

Average Paid Wage by Job Title Subgroup

Q1 Part A

Top Companies by Job Title Subgroup and Average Paid Wage

Q1 Part B

Average Paid Wage Per Year of Data Analysts by State

Avg. Paid Wage Per Year

15,000   115,000

Work State: West Virginia
Avg. Paid Wage Per Year: 109,427

Q1 Part C

Adjusted Average Paid Wage Per Year of Data Analysts by State According to Cost of Living

Avg. Adjusted Wage
11,500   115,000

Work State Abbreviation: WV
Avg. Adjusted Wage: 121,181
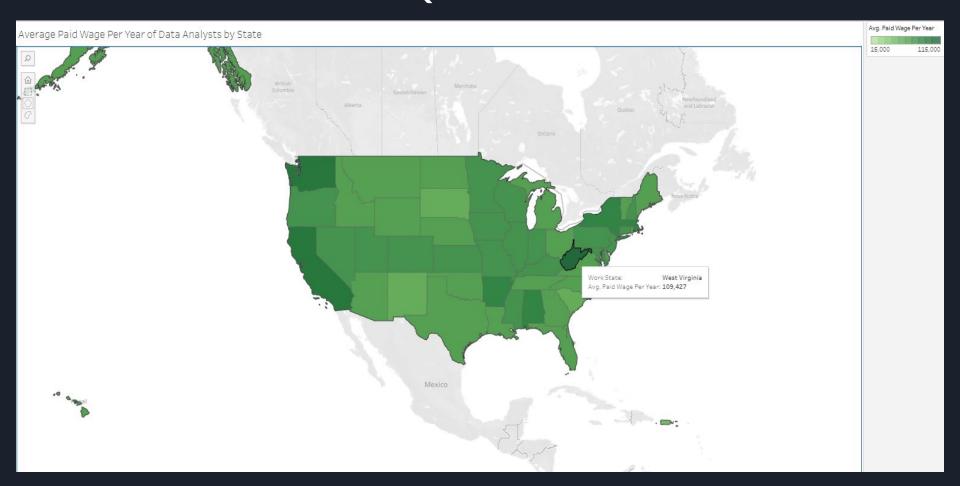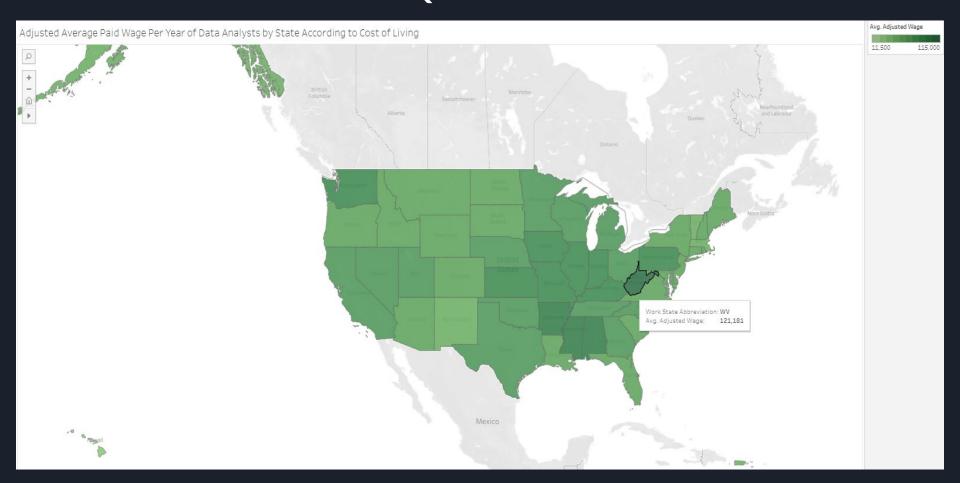
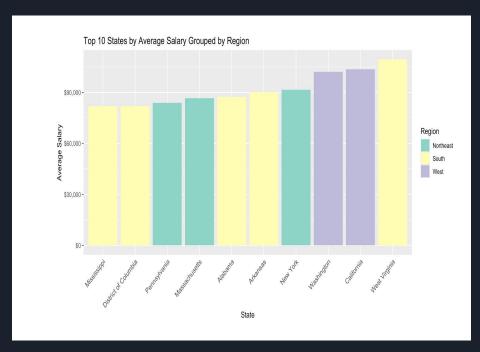# Question #2: What states have the highest paying data-related salaries ?

2a - Differences between job subcategories ?
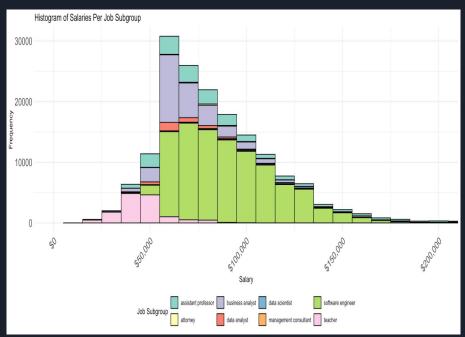
2b - Which companies have the highest salaries of those sub-types ?

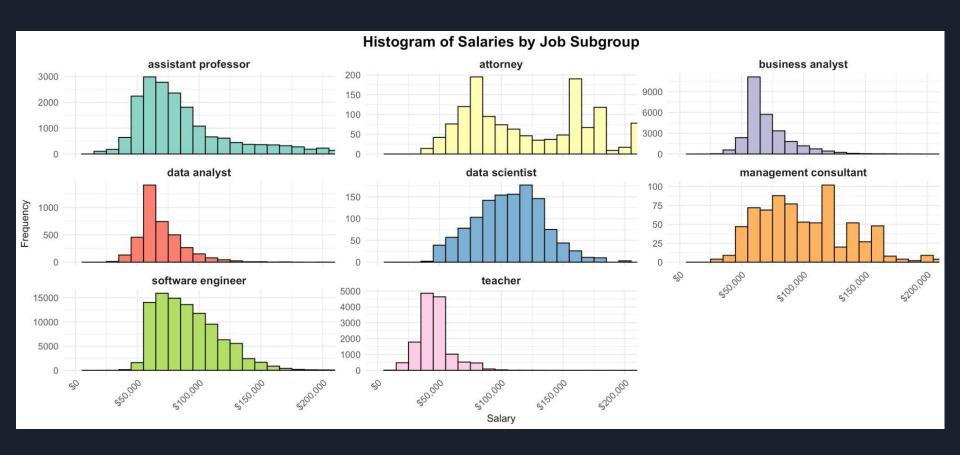2c - Will the answer change if I take the standard of living into account ?
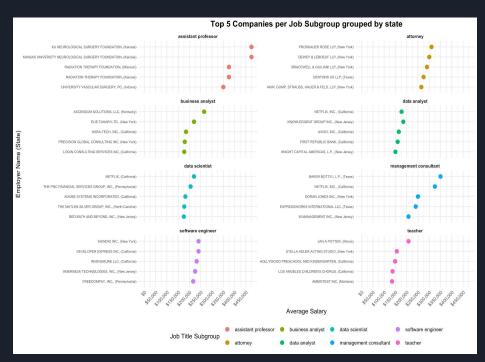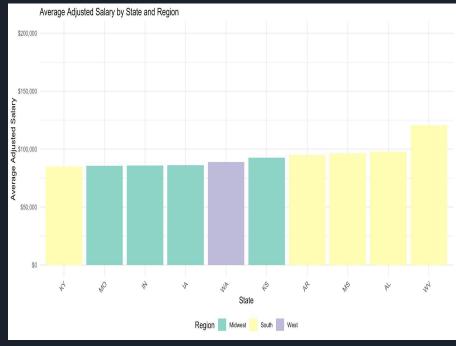
Q2

Q2 Part A

Q2 A, pt II

Histogram of Salaries by Job Subgroup

Q2 Part B

Q2 Part C

Top 5 Companies per Job Subgroup grouped by state

assistant professor
- KU NEUROLOGICAL SURGERY FOUNDATION, (Kansas)
- KANSAS UNIVERSITY NEUROLOGICAL SURGERY FOUNDATION, (Kansas)
- RADIATION THERAPY FOUNDATION, (Missouri)
- RADIATION THERAPY FOUNDATION, (Kansas)
- UNIVERSITY VASCULAR SURGERY, PC, (Indiana)

attorney
- PROSKAUER ROSE LLP, (New York)
- DEWEY & LEBOEUF LLP, (New York)
- BRACEWELL & GIULIANI LLP, (New York)
- DENTONS US LLP, (Texas)
- AKIN, GUMP, STRAUSS, HAUER & FELD, LLP, (New York)

business analyst
- ASCENDUM SOLUTIONS, LLC, (Kentucky)
- ELIE TAHARI LTD., (New York)
- HARA-TECH, INC., (California)
- PRECISION GLOBAL CONSULTING INC, (New York)
- LOGIN CONSULTING SERVICES INC, (California)

data analyst
- NETFLIX, INC., (California)
- KNOWLEDGENT GROUP INC., (New Jersey)
- AVISO, INC., (California)
- FIRST REPUBLIC BANK, (California)
- KNIGHT CAPITAL AMERICAS, L.P., (New Jersey)

data scientist
- NETFLIX, (California)
- THE PNC FINANCIAL SERVICES GROUP, INC., (Pennsylvania)
- ADOBE SYSTEMS INCORPORATED, (California)
- THE MATLEN SILVER GROUP, INC., (North Carolina)
- BED BATH AND BEYOND, INC., (New Jersey)

management consultant
- BAKER BOTTS L.L.P., (Texas)
- NETFLIX, INC., (California)
- DORAN JONES INC., (New Jersey)
- EXPRESSWORKS INTERNATIONAL LLC, (Texas)
- M MANAGEMENT INC., (New Jersey)

software engineer
- SHINDIG INC., (New York)
- DEVELOPER EXPRESS INC, (California)
- INVENSHURE LLC, (California)
- AMERINDIA TECHNOLOGIES, INC., (New Jersey)
- FREEDOMPAY, INC., (Pennsylvania)

teacher
- JAN A POTTER, (Illinois)
- STELLA ADLER ACTING STUDIO, (New York)
- HOLLYWOOD PRESCHOOL AND KINDERGARTEN, (California)
- LOS ANGELES CHILDREN'S CHORUS, (California)
- AMERITEST INC, (Montana)

Employer Name (State)

Average Salary

$0 $50,000 $100,000 $150,000 $200,000 $250,000 $300,000 $350,000 $400,000 $450,000

Job Title Subgroup
- assistant professor
- attorney
- business analyst
- data analyst
- data scientist
- management consultant
- software engineer
- teacher

Average Adjusted Salary by State and Region

Average Adjusted Salary

$200,000
$150,000
$100,000
$50,000
$0

State: KY, MO, IN, IA, WA, KS, AR, MS, AL, WV

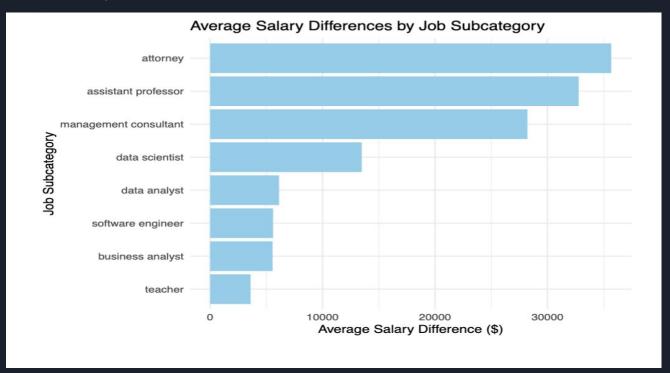**Region**
- Midwest
- South
- West

Question #3: How do offered salaries compare to the prevailing wage ?

3a - Are there job sub-categories that tend to get over- or under- paid ?
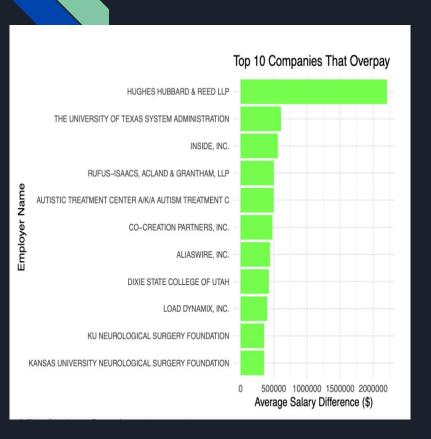
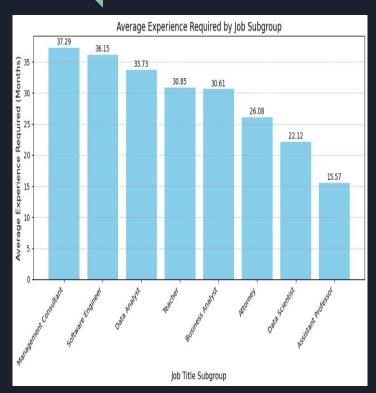3b - Are there companies that tend to over- or under- pay ?

3c - Will the answer change if I take the standard of living into account ?

Are there job sub-categories that tend to get over- or under- paid ?

Average Salary Differences by Job Subcategory

3b - Are there companies that tend to over- or under- pay ?

Top 10 Companies That Overpay

Top 10 Companies That Underpay

3c - Will the answer change if I take the standard of living into account ?
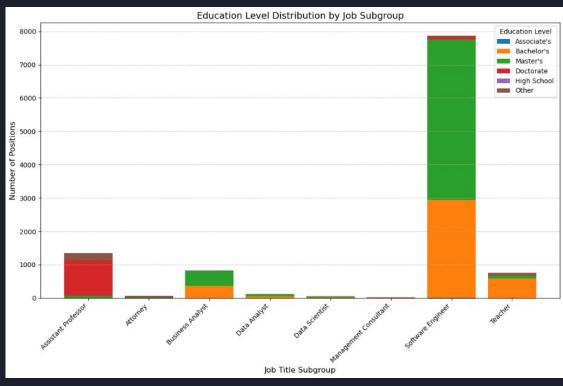
The "Average Adjusted Salary Difference by Job Subcategory (Considering COLI)" table is sufficient to evaluate how COLI influences over- or underpayment. Adjusted salary differences already incorporate cost-of-living factors, making this data appropriate.
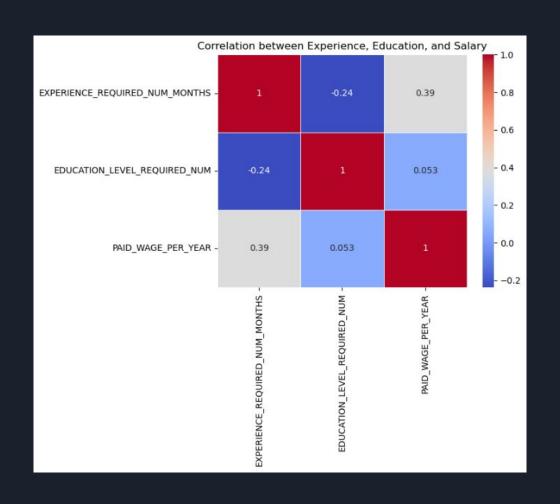
Question #4: How do salaries differ per geographic region ? Are there specific data-related roles found in these regions ?

Composition of Salaries by Region for Each Job Subtype

Question #5: What specific sub-types require higher experience and education, is this correlated with salary ?

Correlation between Experience, Education, and Salary

Question #6: What salaries are "normal" when looking at a distribution of salaries. Is there any trends or patterns found within this salary distribution.

6a - What are the mean, median, and mode of the salary distribution, and how do they differ from each other?

6b - Is the distribution of salaries skewed (positively or negatively), and what does the standard deviation tell us about the spread?

6c - Are there identifiable trends or salary bands that are more common for specific groups, such as by industry, experience level, or geographic region?
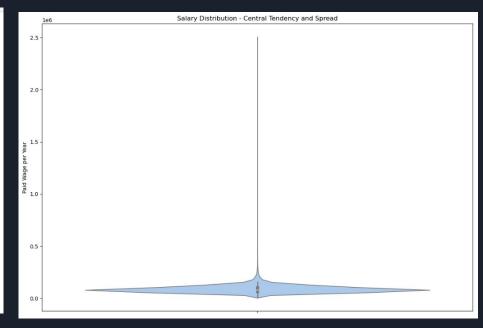
Question 6a

```
Salary Summary:
count     1.672100e+05
mean      8.553035e+04
std       3.872227e+04
min       1.050000e+04
25%       6.300000e+04
50%       7.860150e+04
75%       1.000060e+05
max       2.500000e+06
Name: PAID_WAGE_PER_YEAR, dtype: float64
Mean Salary: 85530.34612535135
Median Salary: 78601.5
Mode Salary: 60000.0
```
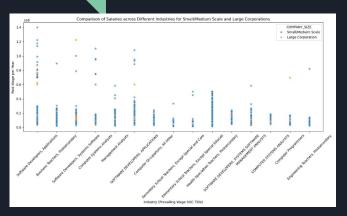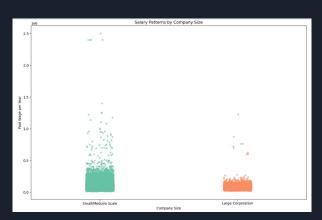
Question 6b

Salary Distribution - Central Tendency and Spread
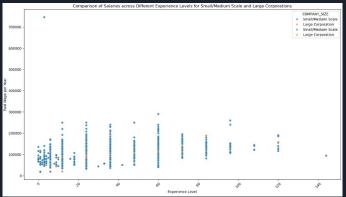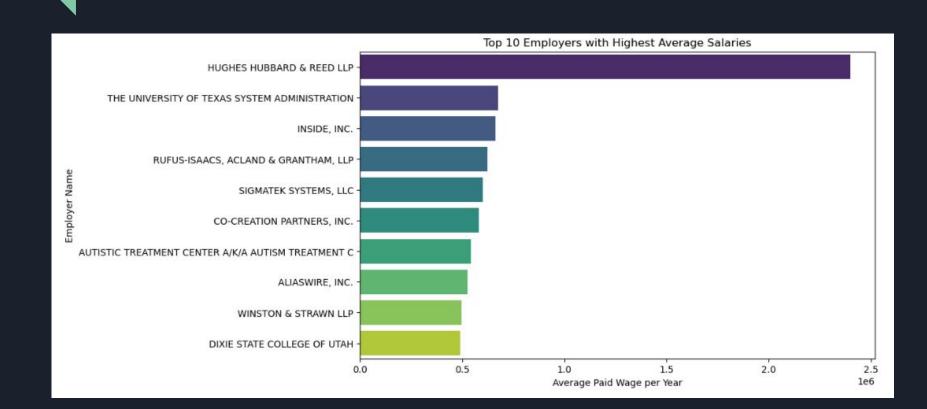
Question 6c

Question #7: How does company size or industry influence salary levels for data-related roles?

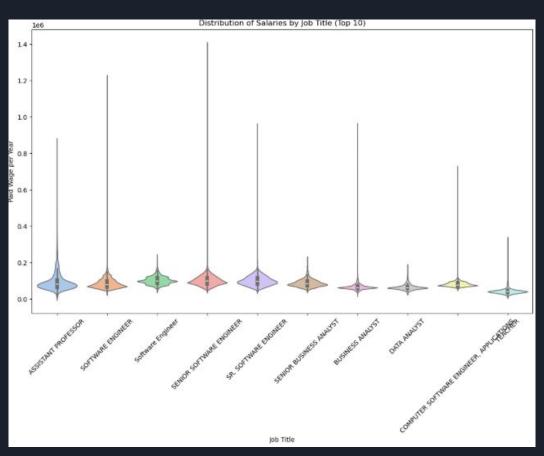7a - Do larger companies or certain industries tend to offer higher salaries for similar roles?

7b - Is there a significant salary difference within industries (e.g., Tech vs. Healthcare) for the same job title?

7c - Do small startups or large corporations offer better salary growth opportunities over time?

Question 7a

Top 10 Employers with Highest Average Salaries

Question 7b

Question 7c

Comparison of Salaries at Startups vs Large Corporations

Network Graph of Employers, Work States, and Salary Levels