Maria Estrada

Professor Blanche Cohen

Computer Science Principles

December 9, 2022

Data Analysis Final Project

**Data Tools**

Data tools section goes over storing data, computing basic statistics, and finding patterns in data. CVS files, spreadsheets, and databases stare data. Databases are the best for storing data (such as SQL) because it can be easily accessed, updated, queried, and deleted. To best compute data graphs need to be made and trends need to be found. From finding these trends predictions can made.

**Big Data**

There is tremendous amount of data being always created. Traditional ways of sorting and processing data is no longer efficient. These big data sets are used in scientific research, digital libraries, and medical records. With such large amounts of data there are challenges with storage, processing and are they ethical.

**Bias in Machine Learning**

Machine learning algorithms learn using reinforcement learning, unsupervised machine leaning, and supervised machine learning. These algorithms are used in many applications such as in criminal justice, hiring and language translation. Another big way machine learning is used is facial recognition which has significant drawbacks like bias, inaccuracy, injustice, mass surveillance, racial profiling, violation of human rights and transparency.

**Unit Test**

The unit tests your understanding of all the units using real world examples. The data gives you a situation, a data set and question based on the data. The questions are based on making predictions, knowing what you can find and what you cant find.

**Step 1: Choose data set and extract info needed**

|  |  |
| --- | --- |
| Weld County, Colorado | https://data.census.gov/table?q=Weld+County,+Colorado+income&tid=ACSST1Y2021.S1901 |
| Median Income | $85,290 |
| Mean Income | $103,262 |

|  |  |
| --- | --- |
| Boulder County, Colorado | https://data.census.gov/table?q=boulder+County,+Colorado+income |
| Median Income | $90,168 |
| Mean Income | $131,470 |

|  |  |
| --- | --- |
| Denver County, Colorado | https://data.census.gov/table?q=denver+County,+Colorado+income |
| Median Income | $81,630 |
| Mean Income | $115,122 |

|  |  |
| --- | --- |
| El Paso County (Colorado Springs), Colorado | https://data.census.gov/table?q=el+paso+County,+Colorado+income |
| Median Income | $79,427 |
| Mean Income | $100,800 |

|  |  |
| --- | --- |
| Larimer County (Fort Collins), Colorado | https://data.census.gov/table?q=larimer+County,+Colorado+income |
| Median Income | $78,109 |
| Mean Income | $105,640 |

|  |  |
| --- | --- |
| Colorado (State) | https://data.census.gov/table?q=colorado+income |
| Median Income | $82,254 |
| Mean Income | $110,650 |

**Step 2: Formulate questions to be investigated**

* Which county has the highest and lowest mean income?
* What counties make above the Colorado mean income?
* Which county has the highest and lowest median income?
* What counties make above the Colorado median income?

**Step 3: Analyze data**

**Step 4: Create Presentation of the process and results**

I wanted to determine what county in Colorado had the highest and lowest income compared to the state average. I choose specific counties based on the cities they contain except for weld county which is where I live. Using [Census Bureau Data](https://data.census.gov/) I found the median and mean income of very county that I choose and the overall state. I put the information found in the census bureau in an excel sheet and created graphs from it which helped me find my answers. As shown below…

* Which county has the highest and lowest mean income?
  + The highest mean income is Boulder County with $131,470. The lowest mean income is El Paso County where Colorado Spring is located with $100,800.
* What counties make above the Colorado mean income?
  + Boulder and Denver County are the only counties that make above the Colorado mean income.
* Which county has the highest and lowest median income?
  + He highest median income is boulder county with $90,168. The lowest median income is Larimer County where Fort Collins is located with $78,109.
* What counties make above the Colorado median income?
  + Weld and Boulder County are both above the Colorado median income.

This means that the overall highest mean and median income is in Boulder County which contains cities like Boulder and Louisville, Colorado. Following this the next mean is Denver County which contains Denver, Colorado which means that they have the second highest average income. But the next median income is Weld County whose biggest city is Greely, Colorado. The Weld County graph is left skewed meaning that more people make higher salaries. The county with the lowest mean income is El Paso County whose biggest city is Colorado Springs which means they have the lowest average. The lowest median income is in Larimer County which contains Fort Collins, Colorado. Which means the graph is skewed right meaning that most people make a lower income.