Robert Pimentel

BANA 698

February 24, 2025

<u>**Frustration Report for Week 3**</u>

<u>**Challenges Faced**</u>

During Week 3, we were tasked with uploading our selected tables to SQLite3 to form a

Database and preprocess/clean the selected indictor chosen. However, for week 3, my biggest

challenged faced was determining the best approach for cleaning the social indicators table.

Majority of the data inside the table was not missing and there were some indicators that needed

a little cleaning. But the most difficult ones were the following.

Multidim_PovertyHeadcountRatio_UNDP, Multidim_PovertyHeadcountRatio_WorldBank,

Poverty_HeadcountRatio_NationalPovertyLines_Percent,

Poverty_HeadcountRatio_SocietalPovertyLines_Percent

I also struggled initially choosing the right csv file to use for the preprocessing/cleaning.

Using the raw excel file instead of the one from the database resulted in the having a bunch of

indicators assigned to one value and cleaning that across the file would result in serious errors.

<u>**Strategies Used to Overcome Obstacles**</u>

The main strategies used to overcome this for this week was a lot of communication with

my team members. When cleaning the social indicators table, I was constantly double checking

with my teammates to make sure we were on the same page and as well as using correct

strategies with normalizing the data. I also referred to videos on how to export a csv from the DB

browser extension. I also got assistance from Ashely Cortez to download the DB browser as it

made navigating the sqlite a lot easier with the extension's user-friendly browser.

<u>Reflections and Lessons Learned</u>

If I were to start this project all over again, I would watch more videos to help me with

sql. Although our sql class was very brief and taught us the basic there is still much more I do

not know about it. I would also like to learn how to use sql through the command prompt

through my computer. I would also definitely look up more ways and strategies on how to

clean/preprocess data because there are more than just the ways that we have learn in class.