Mini-Project 2: Where should be our next campus? (Tables in Excel and Visualization)

Instructions:
SMU currently operates its main campus in Dallas, TX and has two smaller campuses in Plan, TX and Taos, NM. SMU is planning to open a new major campus. Where in the US should this campus be located? The goal is to supplement the strategic decision with an understanding of population growth in different areas of the US. Remember, building a university campus is a long-term decision.

This homework is a group exercise (three students). We will work on the homework in class, and you will finish it at home. The deliverables are a complete report using this template and an Excel workbook uploaded on Canvas. Grading will be based on the report. The report needs to contain the important parts of the workbook pasted into the appropriate places.

Here are some important points:

- A description and discussion of the numbers and charts are essential.
- Highlight your recommendation based on the data.
- You need to convince the reader why your findings are important.
- Make sure your report looks professional (formatting, references, organization).

Executive Summary: [½ page description of the project highlights. What are the key results?]

Table of Contents

1. Frame the problem ..............................................................................................................................2
2. Collect the needed data .......................................................................................................................2
   Data Sources ........................................................................................................................................2
3. Prepare and explore the data ...............................................................................................................2
   Clean and understand the data ..........................................................................................................2
   Format as Table ...................................................................................................................................2
   Pivot Table ..........................................................................................................................................2
   Visualization .........................................................................................................................................2
4. Models and Algorithms .....................................................................................................................2
5. Communicate the results and/or implement a data-driven product ..................................................3
6. Evaluate the value of the project .......................................................................................................3
1. **Frame the problem**
   [What it the question you can answer? Why are they important?]

   Here are some example questions:
   
   - What counties in the US enjoy the highest growth rates?
   - What counties have the lowest growth rates?
   - Do you see patterns?
   - What does the growth rate depend on? Do some research about a few counties/states with high and low growth rates.
   - How would this data influence the decision of locating a university campus there?

2. **Collect the needed data**
   [Data source, data quality and reliability.]

   **Data Sources**

   Use this data source (the whole US):
   
   - [https://www.census.gov/data/tables/2017/demo/popest/total-cities-and-towns.html#ds](https://www.census.gov/data/tables/2017/demo/popest/total-cities-and-towns.html#ds)

3. **Prepare and explore the data**
   [Clean and connect the data.]

   **Clean and understand the data**

   Are there missing data? Does the data make sense?

   The data seems to be a mixture of states, cities and counties. We are interested in counties and states. Clean the data.

   **Hint:** There is a Glossary link on the top of the data page. Maybe that helps!

   Explain how you cleaned the data.

   **Format as a Data Table**

   Use conditional formatting, slicers to help with exploring the data.

   Calculate how much each county has grown. What is the advantage/disadvantage of reporting absolute growth in number of people and growth in percentage?

   **Pivot Table**

   Use a Pivot table to aggregate to states. Make sure that the Pivot table only uses data from counties (check the population of several states manually). You might need a filter or slicer.

   **Visualization**

   Use appropriate visualization.

4. **Models and Algorithms**
   [What type of problem do we have? Forecast values, make yes/no decisions, compare data?]
How will the county population change over the next 10, 25 or 50 years? Can you use the data to say something about this? What are your assumptions?

5. Communicate the results and/or implement a data-driven product
[Prepare report, visualizations, real-time dash boards. Implement decision support tools or incorporate algorithms in apps and web sites.]

What did you learn from the data? How can the data be used?

6. Evaluate the value of the project
[Does the project answer the initial questions?]

Why is this project valuable to the university leaders? What else should they look at?

References

Do not forget references!