Kalle Reaves

5/29/25

Senior Capstone project section 01f

**1. Project Planning & Management**

1.1 Define project objectives and scope  
1.2 Identify data sources (Kaggle, Zillow, Census, BEA, etc.)  
1.3 Develop project timeline  
1.6 Schedule regular progress checkpoints

**2. Data Collection**

2.1 Research housing price datasets  
2.2 Download Zillow historical housing price data  
2.3 Collect inflation rate data from BEA  
2.4 Gather interest rate data from Federal Reserve  
2.5 Collect demographic/economic data from U.S. Census  
2.6 Acquire property-level data with location and features  
2.7 Document metadata and data sources

**3. Data Storage & Organization**

3.1 Design SQL database schema  
3.2 Create SQL tables for each dataset  
3.3 Populate tables with collected data  
3.4 Normalize and index SQL tables  
3.5 Back up SQL database

**4. Data Cleaning & Preprocessing**

4.1 Handle missing values  
4.2 Remove duplicates  
4.3 Standardize data formats   
4.4 Convert SQL tables to pandas DataFrames  
4.5 Perform exploratory data analysis (EDA)  
4.6 Merge data across sources by ZIP code/state

**5. Feature Engineering**

5.1 Create new variables (e.g., price per square foot)  
5.2 Encode categorical variables  
5.3 Create dummy variables for amenities  
5.4 Normalize/standardize numerical features  
5.5 Generate time-based features for time series analysis

**6. Statistical & Predictive Analysis**

6.1 Conduct correlation analysis  
6.2 Perform linear regression to assess price factors  
6.3 Run multivariate regression analysis  
6.4 Conduct time series analysis   
6.5 Test models for inflation-adjusted pricing  
6.6 Evaluate model accuracy

**7. Geospatial & Visual Analysis**

7.1 Map housing prices by ZIP code and state  
7.2 Create heat maps of housing cost increases  
7.3 Visualize inflation-adjusted housing price trends  
7.4 Plot bubble charts of most impactful features  
7.5 Generate time series graphs of price growth  
7.6 Build dashboards using Tableau or similar tools

**8. Interpretation & Reporting**

8.1 Summarize key findings from regression models  
8.2 Compare price increases to inflation  
8.3 Interpret geospatial trends  
8.4 Predict future housing affordability  
8.5 Write detailed project report  
8.6 Include graphs and visualizations in report

**9. Deliverables & Submission**

9.1 Prepare final Python/Jupyter Notebooks  
9.2 Prepare and document SQL source code  
9.3 Export all visualizations (PNG/interactive)  
9.4 Compile and organize results of models  
9.5 Submit full project package (code, visuals, report)