**Data Analytics Capstone Topic Approval Form**

*The guidelines for the Data Analytics Graduate Capstone course require a student to demonstrate the "application of academic and professional abilities developed as a graduate student. The capstone challenges students to integrate skills and knowledge from several program domains into one project." As a result, it is highly recommended that your topic* ***should resolve a current or perceived business problem.*** *As stated, you want to exemplify what you have learned in your MSDA program to showcase your skills. Remember, your research topic exemplifies scholarship and research at the highest level and is significant and helpful to potential employers in identifying your abilities. Your research will always be something you can look back on with pride. Finally, it is recommended for students to use publicly available data sets for transparency and external validity.*

*The purpose of this document is to help you clearly state the research question you will be exploring in your capstone project, your project's scope, and your timeline in order to ensure that these align with your degree emphasis. Without clearly defining each of these areas, you will not have a complete and realistic overview of your project, and it cannot be accurately assessed whether your project will be acceptable for the purposes of these courses. Of course, if this a project that you have already completed at work or elsewhere, this should be easy to fill in! Many students do use a project they have already completed in the past year or two. In that case, you will write the proposal as if the project had not been done yet, and when you report on your project, you will use your complete after-implementation report. If you have not yet done your project, this document can help make sure the scope is within the acceptable range for this capstone. A course instructor will approve this form before submitting this task for evaluation. The task will not be evaluated without a course instructor's signature. The course instructor may ask for additional information before approving the form.*

***Please only submit a Topic Approval Form that has been signed by a course instructor for evaluation.***

**Student Name:** Hien Ta

**Student ID:** 000600660

**Capstone Project Name:** Forecasting Real Estate Market Trends in 2024 Using Time Series Analysis

**Project Topic**: Analyzing and forecasting real estate market trends in 2024 Dallas, Texas, using time series modeling techniques.

**This project does not involve human subjects research and is exempt from WGU IRB review.**

**Research Question:** *Summarize one question or decision you will address by collecting…*What key indicators can predict the real estate price trends in the Dallas market for 2024, and how accurately can time series modeling forecast these trends?

**Hypothesis**: **Null hypothesis**-. There is no significant relationship between the identified indicators and real estate prices. Alternate Hypothesis: The identified indicators significantly affect real estate prices and can be used to forecast future trends.

**Context:** *Explain why the situation or question would benefit from data analysis in less than 500 words.*The real estate market is dynamic and influenced by various economic, demographic, and policy factors. An accurate forecast model can help investors, policymakers, and stakeholders make informed decisions.

**Data:** *Identify data you will need to collect relevant to the situation or question.*I will be utilizing the USA Real Estate Dataset from Kaggle, which includes real estate listings with details such as listing prices, bedrooms, bathrooms, square footage, and other features across the United States, with the ability to focus specifically on Dallas, Texas.

*If an existing data set will be used, describe the data set.* The USA Real Estate Dataset on Kaggle is a collection of over 2.2 million real estate listings across the United States. The dataset is organized by state and zip code, which allows for extracting and analyzing information specific to a particular region, such as Dallas, Texas. The data includes listing prices, number of bedrooms, number of bathrooms, square footage, and other relevant features that can be used to comprehensively analyze real estate trends.

*Explain who owns the data and why you are allowed to use this data for your capstone project.* The dataset is owned by Ahmed Shahriar Sakib and made available on Kaggle. Kaggle datasets are made available for educational and research purposes.

*Note: If you are using restricted information, please have the Third-Party Authorization Form signed by an authorized agent on behalf of the data owner. The data owner's legal name is required on the form.*

**Data Gathering:** *Describe the data-gathering methodology you will use to collect data.*

Ahmed Shahriar Sakib on Kaggle has already collected and made available this dataset. As a publicly available dataset, it is assumed that it can be used for academic and research purposes, provided the terms of service of Kaggle are met.

**Data Analytics Tools and Techniques**: *Identify the appropriate data-analysis technique to analyze this data.*

To analyze this dataset, time series modeling techniques such as ARIMA, SARIMA, or LSTM networks will be applied to understand and forecast real estate price trends, considering factors such as seasonality, economic cycles, and policy changes.

**Justification of Tools/Techniques:** *Explain why the data-analysis technique you chose is appropriate for analyzing the data collected*.

The chosen time series analysis tools are suitable for the dataset as they are designed to handle sequential and indexed data over time, which is characteristic of real estate price data. The tools will be selected based on the data's behavior, trend, and seasonal patterns observed in the preliminary analysis.

**Project Outcomes**: *List the key anticipated project outcomes and deliverables in less than 500 words*.

The anticipated outcome is a detailed model capable of forecasting real estate prices in Dallas, supported by an analysis report that measures the model's predictive power and accuracy.

**Projected Project End Date**: 5/15/2024

**Sources**:

**(**Ahmed Shahriar Sakib, 2024**)https://www.kaggle.com/datasets/ahmedshahriarsakib/usa-real-estate-dataset**

**Course Instructor Signature/Date:**

The research is exempt from an IRB Review.

An IRB approval is in place (provide proof in appendix B).

Course Instructor's Approval Status: Approved

Date: Click here to enter a date.

Reviewed by:

Comments: Click here to enter text.