**Data Analytics Capstone Topic Approval Form**

*The capstone challenges students to integrate skills and knowledge from several program domains into one project. The guidelines for this capstone course require you to demonstrate the application of academic and professional abilities developed as an undergraduate student in the BSDA program. It is highly recommended that the topic of your capstone be about resolving a current or perceived business problem. Your research topic should exemplify scholarship and research at the highest level and should be significant enough that it would help potential employers identify your abilities. It is also recommended that you use publicly available datasets for transparency and external validity.*

*This document is designed to help you clearly state the research question you will be exploring in your capstone project, the scope of your project, and your timeline in order to ensure that all of 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 this capstone course.*

*If your project is one you have already completed at work or elsewhere, this document should be easy to complete. Many students do use a project they have already completed in the past. In that case, you will write the proposal as if the project has not been completed yet, and when you report on your project, you will use your complete after-implementation report. If you have not yet completed your project, this document can help ensure the scope is within the acceptable range for this capstone. An instructor must approve this form before you submit this task for evaluation. The task will not be evaluated without an instructor’s signature. The instructor may ask for additional information before approving this form.*

*Before submitting this form for approval, please remove all italicized directions in the form.*

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

**Capstone Topic Approval Form**

The purpose of this document is to help you clearly explain your capstone topic, project scope, and timeline. Identify each of the following areas so you will have a complete and realistic overview of your project. Your course instructor cannot approve your project topic without this information*.*

*Note: You must fill out and submit this form. Space within each section will expand as needed.*

*Any costs associated with the development of the data analytics solution will be the responsibility of the student.*

**Student Name:** Jonathan Mitchell

**Student ID:** 011300756

**Capstone Project Name:** Steam Game Insights: Market Trends from Metadata Analysis

**Project Topic:** Steam Game Insights: Analyzing Steam Game Metadata for Market Trends

**Research Question:** *Summarize one question or decision you will answer by collecting and analyzing a set of data.*What factors—such as genre, price, and release timing—correlate with higher owner estimates for Steam games?

**Hypothesis:** Steam games in genres like Action/RPG, priced in the mid-tier, and released during peak periods (e.g., holidays or early year) have higher estimated owner counts.

**Context:** *Explain in 500 words or less why the situation or question would benefit from data analysis.*As the indie game development industry grows, understanding market dynamics on platforms like Steam is essential for game developers. Steam is the largest digital distribution platform for PC games, and its sales trends provide insights into what kinds of games gain traction. By analyzing a comprehensive dataset of Steam games, including features such as genre, price, and release date, this project seeks to determine patterns that correlate with game popularity (measured by owner estimates). These insights will support more informed design and release strategies for developers. For me personally, this project doubles as market research to shape the development direction and marketing of future game projects.

**Data:** *Identify data you will need to collect that are relevant to the situation or question.* **Kaggle “Steam Store Games” dataset** (∼27 k games with genre, price, owner estimates) [kaggle.com+9kaggle.com+9kaggle.com+9](https://www.kaggle.com/datasets/nikdavis/steam-store-games?utm_source=chatgpt.com)

*If an existing dataset will be used, describe the dataset. The dataset was collected by Kaggle user Justen Bennett and contains metadata for thousands of games listed on the Steam store. It includes price, genre, estimated number of owners, release date, and developer/publisher information.*

*Explain who owns the data and why you are allowed to use the data for your capstone project.* The dataset is publicly available on Kaggle under a CC0 (public domain) license. The data is compiled from publicly available information on the Steam storefront and is not proprietary.

*Note: If you are using restricted information, please have the “Authorization to Use Restricted Information” 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.*Download directly via Kaggle API or web interface. Load into pandas DataFrames for analysis.

**Data Analytics Tools and Techniques:** *Identify the appropriate data-analysis technique you will use to analyze the data.* Python (pandas, matplotlib/seaborn for EDA, scikit-learn for modeling), Exploratory data analysis, Regression modeling, Correlation analysis, Visual storytelling through graphs and charts

**Justification of Tools/Techniques:** *Explain why the data-analysis technique you chose is an appropriate technique to analyze the data collected*. Python is well-suited for data cleaning, exploration, and statistical analysis. The selected techniques allow identification of patterns and relationships between game characteristics and ownership levels, which is the core objective of this project.

**Application Type, if applicable (select one):**

mobile

web

stand-alone

**Programming/Development Language(s), if applicable:** Python 3.x

**Operating System(s)/Platform(s), if applicable:** Windows

**Database Management System, if applicable:** None — data stored in flat CSV format and manipulated using pandas.

**Project Outcomes:** *List the key anticipated project outcomes and deliverables in fewer than 500 words*. Cleaned dataset ready for analysis Visualizations showing distribution of genres, price tiers, and release dates Correlation matrix of game features vs. owner estimates Regression or classification model predicting ownership level Market insight report that could guide future indie game development efforts

**Projected Project End Date:** 7/25/2025

**Sources:** Kaggle dataset: <https://www.kaggle.com/datasets/justinbenedict/steam-store-games> Python documentation, matplotlib, seaborn, scikit-learn documentation

**Human Subjects or Proprietary Information**

Does your project involve the potential use of human subjects? (Y/N): N

Does your project involve the potential use of proprietary company information? (Y/N): N

**STUDENT SIGNATURE**

**­­­­­­­­­­­­­­­­­­­­­­­­­­­­­\_Jonathan Mitchell\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**By signing and submitting this form, you acknowledge** that any cost associated with the development and execution of your data analytics solution will be your (the student) responsibility.

**TO BE COMPLETED BY AN INSTRUCTOR**

**The capstone topic is approved by an instructor.**

**INSTRUCTOR’S NAME AND SIGNATURE:**

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**INSTRUCTOR APPROVAL DATE:**

**Project Compliance with IRB (Y/N):**