

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: Cooper Hepworth

Student ID: 011280038

Capstone Project Name: "Constructor vs Racer: An Analysis of Formula 1 Racing"

Project Topic: Exploring Formula 1 race data to observe impact of constructor vs racer

Research Question: *Summarize one question or decision you will answer by collecting and analyzing a set of data.* Is the constructor or racer a better predictor for lap time on a given track.

Hypothesis: The driver of the car is a better predictor for lap speed than the constructor on a given circuit.

Context: *Explain in 500 words or less why the situation or question would benefit from data analysis.* There are multiple factors that contribute to the success of a Formula 1 racer. The driver must utilize incredible skill between the wheel to complete the track quickly. This involves reaction time, finding and maintaining the optimal driving line at the optimal speed, and keeping control of the vehicle through winding turns. Behind the scenes a team of mechanics works to ensure that the racer holds together and performs as optimally as possible. The engine of a Formula 1 racer requires careful maintenance. Components such as pistons, sparkplugs, and tires have to be maintained and replaced frequently. This ensures the vehicle's performance as well as the driver's safety. The constructor is the brains behind the design of the racer. They must take careful consideration when designing the engines and layout of the cars. What kind of materials will be involved in the construction of the engine? How much displacement will the engine have? What will be the piston configuration? Advanced mechanical engineering and mathematics go into the designing and development of a Formula 1 racer. What is observed on race day is only the tip of the iceberg. Formula 1 cars are expensive. Formula 1 teams are given a price limit per season of \$135 million. The car itself can cost around \$15 million. Teams are allowed 3 engines per season; each engine can cost up to \$10 million. Factor in the cost the chassis, gearbox, and tires used and replaced throughout the season and you start to understand the \$135 million price point. (see the following article for more information: [How much does an F1 car cost in 2024? Key parts, history of most expensive Formula 1 cars](#)). A team's success can be defined by their ability to estimate the best way to stay within their budget. Data could help with the following scenario. A manufacturing company sees a lucrative advertising opportunity in Formula 1 racing. They decide to enter the competition. A portion of the marketing budget is given to you to build the racing team. You



have to decide how much you want to invest in acquiring a skilled driver and what kind of car you want to build for the race. Your goal is to maximize performance on the track by getting better lap times. The car wearing your company's logo will be a more effective advertisement the better it performs. Do you allocate more of your money towards R&D or finding a skilled driver? This is where data comes in. The object of this study seeks to observe which has a bigger impact on performance via lap time, the driver or the constructor. This study will hopefully lead to a data-driven decision being made.

Data: *Identify data you will need to collect that are relevant to the situation or question.* Formula 1 race data best lap time, finishing position, and overall race time, organized by drivers, constructors, and races.

If an existing dataset will be used, describe the dataset.

<https://ergast.com/mrd/db/#csv> will be used. This dataset contains several .csv files containing data on races, drivers, constructors, and results. It contains all of the data needed for our analysis.

Explain who owns the data and why you are allowed to use the data for your capstone project. Ergast owns the dataset. The site provided states that "Attribution is optional but appreciated when addressing a technical audience. A reference to "Ergast" with a link back to this site is sufficient."

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.* The data will be downloaded as .csv files provided by the site. The data will be static for this report.

Data Analytics Tools and Techniques: *Identify the appropriate data-analysis technique you will use to analyze the data.* The tools used will be python with the pandas package for manipulation and transformation of the data. Other packages such as Matplotlib, Seaborn, and SciKitLearn will also likely be used. The statistical test that will be performed is a chi-squared test of independence.

Justification of Tools/Techniques: *Explain why the data-analysis technique you chose is an appropriate technique to analyze the data collected.* The data is already in a well-structured format and will be relatively easy to work with. Once the data is organized in a proper frame, the next task will be to perform a chi-square test of independence on the data.



Application Type, if applicable (select one):

- ☐ mobile
☐ web
☒ stand-alone

Programming/Development Language(s), if applicable: Python

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

Database Management System, if applicable: None

Project Outcomes: *List the key anticipated project outcomes and deliverables in fewer than 500 words.* The goal will be produce a chi-square statistic testing the driver and constructor in predicting overall track time, best lap time, best lap speed, and finishing position. The test will provide charts showing differences in data.

Projected Project End Date: 12/31/2024

Sources: <https://ergast.com/mrd/db/#csv>

Human Subjects or Proprietary Information

Does your project involve the potential use of human subjects? (Y/N): No

Does your project involve the potential use of proprietary company information? (Y/N): No

STUDENT SIGNATURE

Cooper Sylvester Hepworth

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.

COURSE INSTRUCTOR SIGNATURE:



Jim Ashe, Ph.D. Mathematics

COURSE INSTRUCTOR APPROVAL DATE:

12/6/2024



WESTERN GOVERNORS UNIVERSITY®

Project Compliance with IRB (Y/N): Y

