Data Management and Data Analytics Capstone Topic Approval Form

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.

Student Name: [redacted]

Student ID: [redacted]

Capstone Project Name: Analysis of Growth Performance of Actively and Passively Managed Retirement Portfolios

Project Topic: This project will analyze the overall growth of a single-fund retirement portfolio comparing actively managed "Target Date Retirement" funds and passively managed index funds (Exchange Traded Funds).

Research Question: When comparing historical data of a group of Target Date Retirement funds against a group of index fund Exchange Traded Funds (ETF), is there a noticeable difference in theoretical returns. When conducting simulations of contributions to these funds, do the practical results (simulation) reinforce the theoretical results?

Hypothesis: My hypothesis is index fund ETFs will generally outperform Target Date Retirement funds in both our theoretical returns test as well as our simulated returns test.

Context: Using tax-advantaged retirement accounts to invest in "Target Date Retirement Funds" or index fund ETFs can both be considered viable "hands-off" approaches when it comes to saving for retirement. The "hands-off" and "set it and forget it" philosophies entail contributing to a fund (or fund portfolio) for a consistent long-term basis. Using this approach, there is typically no selling of securities until it is time for an overall portfolio rebalance as once nears retirement. This practice applies the concept of having "time in the market" rather than "timing the market", under the assumption that, generally, the value of stocks will increase over time. If there is a significant difference in performance in a certain group of funds compared to the other, it may be worth exploring options within that group of funds in order to maximize potential returns over time. For compairing best performance we will consider total returns or growth over a period of time. For this experiment, we will not adjust returns based on expense ratios of funds. Note that we will not be comparing all available fund options for all available dates, nor will we be exploring all possibilities of mixtures of asset allcoations; rather this experiement analyzes single fund portfolio performance over a defined period of time.

Data: I will be utilizing Python's yfinance library to programmatically retrieve data from the Yahoo Finance Application Programming Interface (API). This data is free to use for research and educational purposes as long as uses are compliant with Terms of Service (ToS), which can be found on the Yahoo Finance website. To keep a simple and consistent sample, we will utilize funds from both groups ("Target Date Retirement Funds" and index fund ETFs) that can be purchased and traded by the same brokerage. For this project, I have chosen Vanguard. A list of Vanguard "Target Date Retirement Funds" can be retrieved from their website. Their ToS allow use for educational purposes as long as the data is not being used for legal or tax advice. Since this project is an experiment and will not be used a legal, tax, or any kind of financial advice, we are compliant with their ToS.

Data Gathering: I will utilize Python's yfinance and Pandas libraries to programmatically retrieve the datasets that we need for this project. All downstream cleansing, transformations, and analytics will also be conducted in Python.

Data Analytics Tools and Techniques:

Tools: I will utlilize Python as the sole programming language for this project. The foundational logic to retrieve, store, transform, and visualize data will be created in custom modules. I will utilize a Jupyter Notebook to import these modules, retrieve the data, conduct analytics, and present visualizations. This application will be as modular as possible and output can be customized by use of a configuration file.

Techniques: My statistical analysis will begin with conducting theoretical and practical returns with the use of simulations. Using our resulting datasets I will apply descripting statistics to each simulated and theoretical output. I will finally utilize an Independent Samples t-test to compare the means of each strategy. I will rely on the p-value to determine if our results are statistically significant.

Justification of Tools/Techniques: Using descriptive statistics is a great way to summarize datasets so that we can understand the spread and central tendency of each dataset. Conducting a t-test is an excellent way to determine whether there is any significance in difference of means. By utilizing a theoretical approach followed by a practical approach, we will be able to either reinforce or challenge the results from each test.

Application Type, if applicable (select one):

- ☐ Mobile
- ☐ Web
- ☒ Stand-alone

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

Operating System(s)/Platform(s), if applicable: This program will be developed on Windows OS; however, it should be platform independent.

Database Management System, if applicable: N/A

Project Outcomes: This project will determine if there is a significant difference in performance between specific actively managed "Target Date Retirement Funds" and passively managed index fund ETFs. All code, analytical results, and presentations will be included in the final submission. This is for educational and informational purposes only and is not to be used as financial guidance or advice.

Projected Project End Date: [redacted] 2023

Sources:

- Stock Market Financial Data:
  - https://pypi.org/project/yfinance/
  - https://legal.yahoo.com/us/en/yahoo/terms/product-atos/apiforydn/index.html
- Vanguard Target Retirement Date Fund List:
  - https://investor.vanguard.com/investment-products/mutual-funds/target-retirement-funds

---

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
[redacted]

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 FILLED BY A COURSE INSTRUCTOR

The capstone topic is approved by a course instructor.

COURSE INSTRUCTOR'S NAME AND SIGNATURE: James R Ashe

COURSE INSTRUCTOR APPROVAL DATE: Tuesday, [redacted] 2023

Project Compliance with IRB (Y/N): Y