

MScFE 600 FINANCIAL DATA

Group Work Project # 1

[See grading rubric here.](#)

Scenario

The financing team is reaching out to your group to provide data and analytics for their new dashboard. This dashboard will help them determine rates and analytics for working with customer inquiries.

The financing team is tasked with 6 different scenarios in which they would lend. These are:

- 1) Money at a fixed rate for an unsecured purchase (e.g. credit card) for an individual.
- 2) Money at a floating rate for a secured purchase (e.g. home or automobile) for an individual.
- 3) Money at a fixed rate for a business for a construction loan.
- 4) Publicly traded Equity (e.g. common stock) – that is, securities lending of a stock.
- 5) Publicly traded bond (e.g. treasury bond, corporate bond) – that is, securities lending of a bond.
- 6) An illiquid security – you choose the security.

Tasks

Recall from Financial Markets, the first set of challenges commonly encountered in finance are:

1. Collateral Related Risks: Financing & Credit
2. Statistical Related Risks: Volatility & Correlation

Step 1: Collateral Related Risks

First, we'll start with collateral related risks. Build a table with 6 rows (1 for each of the 6 scenarios above) and 2 columns: 1 for financing risks and 1 for collateral risks. This will give a total of 12 boxes. Then, in each box, describe either the financing challenges

(scenarios 1, 2, and 3) or the collateral challenges (scenarios 4, 5, 6) that affect each scenario. For example, let's consider Scenario 2, where the financial team is lending money in a variable-rate mortgage to an individual for the purchase of a home. In this case, the collateral is the home itself. One of the challenges is that the home may go down in value – if there is a housing market crash. Another challenge is that the home may be destroyed by a fire. Note that your group is not providing solutions, but just identifying the comprehensive challenges. You may find there are several bullets to include within a single box. You may also find similarities down a column as some of the scenarios face similar, if not the same, risks.

Groups of 2: Pick only 2 of the money lending scenarios (Scenarios 1, 2, 3) and 2 of the security lending scenarios (Scenarios 4, 5, and 6). This applies to all steps!

Step 2. Statistical Related Challenges

Now, create a 2nd table, also with 6 rows (1 for each of the scenarios above) and 2 columns: 1 for volatility and 1 for correlation. Again, this gives a total of 12 boxes. Again, for each box, describe the volatility challenges in the first column, and the correlation challenges in the second box. For example, let's reconsider Scenario 2. One of the challenges is the volatility of the housing market, which affects the value of the collateral, the debt-to-equity ratio, the default risk, etc. One of the challenges in correlation is the correlation of interest rates and housing defaults – as interest rates increase, the borrower has to pay higher amounts each month, making it more difficult to meet the payments.

Again, do not provide solutions but just a thorough list of the issues.

Step 3. Identifying Data

Now, thinking about the challenges, produce a list of the financial data series you want to collect in each category. Do not worry if you cannot find the data. For now, just think about the types of data you want to collect. Data should be identified with several characteristics:

- 1) Data Type: Asset, Accounting, Economic, Ratings, Factor, ...
- 2) Data Processing: Raw prices / yields; Levels / Categories; Returns; Implied Volatilities, ...
- 3) Data Frequency: High-Frequency, Intraday, Daily, Weekly, Quarterly, Annually, ...

- 4) Data Class: Equity, FX, Crypto, Fixed Income, Credit, Commodity, Derivatives, Real Estate, ...
- 5) Data Source: Exchanges, OTC, Brokers, Dealers, Audited Financials, Vendors, Analytics Companies
- 6) Data Variety: Trade Data vs Quote Data; Actual Data vs Estimated Data; Observed Data vs Modeled Data; Adjusted Data vs Unadjusted; Absolute Data vs Relative Data, etc.

For example, let's consider Scenario 5. Suppose you want to collect recovery rates from the credit rating agencies. Step 3 asks you to identify the data characteristics. The following is a list of how you would likely describe credit ratings data:

- 1) Ratings data
- 2) Levels (e.g. AAA, AA, A, BBB, BB, B, C);
- 3) Varying frequency (as ratings are not necessarily updated with regularity)
- 4) Credit data
- 5) Provided by credit vendors like S&P, Moody's, Fitch (e.g. may be available only to subscribers or from certain vendors)
- 6) Modeled Data (with proprietary modeling done by the agencies)

In this step, you may use AI (i.e., ChatGpt) to identify the types of data, and how the data plays a role in addressing the challenge. Remember: you are NOT permitted to copy and paste the text exactly as is. This is considered plagiarism and will be treated as such. You must comply with the [Academic Policy on the use of AI](#).

Step 4. Go get the data

Some data will be easy to get. Other data may be very difficult. Identify a few series for which you can actually locate the data and import into Python.

In Scenario 2, let's imagine we get the income history for the potential borrower (Clearly, this is not public data). Our next idea is to get publicly available financial data. So then we think about default rates for the borrower's credit rating, is FICO score 700.

Be creative in finding sources for your data!

Step 5. Perform exploratory data analysis

Use Python to graph series, distributions, 2-way relationships, and any other meaningful visuals that show insight into the data. Your group will decide both what to graph, and write the code to create them. **Each visual must include 1 sentence that must explicitly describe the insight.**

Step 6. Describe how the data can help to meet the challenge

Suppose you decided that the VIX (volatility index) was a great data set to use. You show that when VIX increases, your security from Scenario 6 became even more illiquid – because you also decided that volume traded of that security was a great data set to use. You can now show that VIX is a good proxy for the illiquidity of that security. Similarly, show how the data you chose and graphed helps to address (perhaps even solve?) the financial challenge.

Submission requirements and format

One team member submits on behalf of the entire group the following items:

1. **1 PDF document*** with individual answers from steps 1, 2, 3 and 6.
 - a. Use the available Report Template and fill out the required information on the first page
2. A **zipped folder** including:
 - a. .ipynb executable Jupyter notebook**
 - b. 1 PDF document **with the output** from the Jupyter notebook. To include the output, RUN the code before downloading the PDF.

** **Use Google Docs to collaborate.** Start by uploading the Report Template provided in the Course Overview. Once your report is completed, click File → Download → PDF Document (.pdf) to obtain the copy for your submission.*

*** **Use Google Colab or GitHub to collaborate** in completing the executable Python program.*

*The PDF file with your report must be uploaded **separately** from the zipped folder that includes any other types of files. This allows Turnitin to generate a similarity report.*

Rubric

Your instructor will evaluate your group submission for GPW1 using the following rubric:

Quantitative Analysis (open-ended questions)	Technical and Non-technical Reports	Writing and Formatting
40 Points	30 Points	20 Points
<p>The group is able to apply results, formulas, and their knowledge of theory to real-life finance scenarios by doing the following:</p> <ul style="list-style-type: none"> • Providing all the necessary information to support their arguments. • Presenting arguments that reflect group discussion and research. • Using authoritative references to support a position and provide updated information • Concluding with practical takeaways for more insightful financial decision-making 	<p>Technical Reports contain 3 parts:</p> <ol style="list-style-type: none"> 1) summary of key results; 2) interpretation of results; and 3) the recommended course of action that can reasonably follow from those results and interpretations. <p>Note: Technical reports will include the technicalities of models, such as names, methods of estimation, parameter values, etc. and exclude generalities about the work done. It should NOT include the names of Python code that was used.</p>	<p>A submission that looks professional should include:</p> <ul style="list-style-type: none"> • The axes labels and scales in graphs. • No significant grammar errors or typos. • Organized, clear structure, and easy to read document. • Proper citations and bibliography using MLA format.
	<p>Non-technical Reports contain 3 parts:</p> <ol style="list-style-type: none"> 1) clear explanation of results; 2) the recommended course of action that follows; and 3) the identification of factors that impact each portfolio. <p>Note: AVOID all references to model names, algorithms, unnecessary details, and focus on the investment decision.</p>	