ALT 2: Data Analytics Project Report

Candidate Names: Matthew Reynolds and Arjun Kunder

Class Group: ComScience

Teacher Name: Mr McEneaney

School Name: St.Josephs CBS

Project Title: Risk Adjusted Commodities Portfolio

Date Submitted: 01/05/2025

Project Link: https://commodities-finance-dashboard.streamlit.app/

1. Problem Statement

The purpose of this investigation is to analyse a dataset to answer and formulate the data into the most effective risk adjusted portfolio. The chosen dataset contains detailed information on commodities prices, over the last 25 years building the most effective risk adjusted rate of return.

The hypothesis being tested is:

A portfolio weighted for sharpe ratio will have a higher risk adjusted return then any single commodity.

2. Description of the Dataset

- Filename: ./commodities_data.csv
- Source: Publicly available Finance dataset
- Number of Records: Approximately 6500+
- Key Fields Used:
- Name – Name of the Commodities
- Prices over the last 25 year – Numerical prices to be analysed

3. Design and Planning

3.1 Programming Language & Tools

Python was chosen for its suitability in data analysis.
Libraries used:

- pandas – for data handling
- streamlit -to create a dashboard
- matplotlib – to create graphs
- numpy - to mathematical calculations

3.2 Data Cleaning Steps

Replace all missing values by averaging the two previous values and using the given value.

4. Implementation

The Python program performs the following actions:

- Reads the `.csv` file containing commodity price data.
- Cleans the data by forward-filling or averaging missing values.
- Calculates daily returns and uses them to compute:
    - Mean daily return
    - Annualized return (multiplied by trading days)
    - Covariance matrix of daily returns
    - Portfolio standard deviation (volatility)
    - **Sharpe Ratio**
- **Optimizes portfolio weights** using `scipy.optimize.minimize` to maximize the Sharpe Ratio.
- Compares the optimized portfolio with the top single-performing commodity (Gold).
- Displays key metrics (Annual Return, Volatility, Sharpe Ratio) and visuals (line graphs, bar charts, bubble charts, pie charts) on an interactive Streamlit dashboard.

🔍 What is the Sharpe Ratio?

The **Sharpe Ratio** is a measure of **risk-adjusted return**, showing how much excess return you receive for the extra volatility you endure. It helps investors compare the return of an investment relative to its risk.

$$\text{Sharpe Ratio} = \frac{R_p - R_f}{\sigma_p}$$

Where:

- R_p = Annualized return of the portfolio
- R_f = Risk-free rate (e.g., a government bond yield)
- σ_p = Annualized portfolio standard deviation (volatility)

A higher Sharpe Ratio implies a better risk-adjusted return.

To find the best mix of commodities, we used scipy.optimize.minimize. The goal is to find weights for each commodity that maximize the Sharpe Ratio, under these constraints:

- The sum of all weights must equal 1.
- All weights must be between 0 and 1.

5. Analysis of Results

5.1 Key Findings

- Highest returning: Nature Gas had the highest return of all the commodities. - Highest volatility: Oil had the highest volatility of all the commodities.
- Highest risk adjusted return: Gold has the highest risk adjusted return of all the commodities.
- Benefits of weighting by sharpe ratio: Your overall portfolio has a mid - high return with mid - low volatility.

5.2 Graphical Output

- The sliding menu shows you what years you want to include in your calculations.
- The menu has a dropdown to choose commodities and display them in a table.

- Then a line graph showing the annual rate of return of each of the selected commodities.
- There is a option for you to select your desired risk free rate

- The table including all the commodities, their sharpe ratio, annual rate of return and the annual volatility.

- The bar chart showing the sharpe ratio broken down by commodity.

-A toggle to either optimise weights or pick your own

- The table and pie chart showing the weights for each commodity for your portfolio

- The performance metrics, comparing your portfolio and gold in both a written section and bar graph.

5.3 Interpretation

The data strongly supports the hypothesis that a portfolio weighted by sharpe ratio will have good returns with lower level of risks than any single commodity.

6. Evaluation and Reflection

6.1 Effectiveness of the Program

The program performed its tasks correctly and returned accurate data and portfolio's either through the optimisation or through personal creation. Visual output improved understanding of differences between our portfolio and each individual commodity.

6.2 Limitations

The dataset does not account for macroeconomic events (e.g. inflation, war, interest rates). Only focused on specific time period (eg. may not reflect all market conditions).

6.3 Improvements for Future Projects

Include correlation analysis or machine learning techniques based around macroeconomic events. Include more consideration for the weighting process. Consider performance efficiency for larger datasets.

7. Appendix: Full Python Code

https://github.com/MasterHacker2008/finance_dashboard.git