



Data Scientist Position

Technical Business Case Problem – Solution Challenge

As part of the recruiting process for the Data Scientist Position we have developed a Technical Business case Problem – Solution Challenge. This challenge is intended to get some insights on the critical thinking and problem solving skills of the candidate.

Please develop a solution or an explanation of the potential solution or elaborate on why the problem could not be solved.

Case: Risk Analysis on Stable Coins in DeFi

Definition of the problem – challenge:

The Blue Swan Data Science team is building a risk framework that will allow the Operations Team to Risk-Adjust the returns of Stable Coin Liquidity Pools in the market so that they can choose the best pools for Yield Farming and other strategies.

One of the key challenges is finding reliable data sources that consolidate pool data and APYs. One of the data sources identified is Coindix.com which consolidates different pool's data from different chains and protocols.

In this challenge you will be part of the Blue Swan Data Science team helping them achieve one of the data pipeline goals which is extracting the data. The end goal of this challenge is to extract data from Coindix and display a classification of the data using either Streamlit or a Jupyter notebook or any other tool available for visualization.

The challenge consists of:

1. Extract data from Coindix.com either the frontend or through their API
2. Select all the Stable Coin Pools in this data
3. Classify them according to the following rules:
 - Pools Class A are pools with a TVL of more than \$50MUSD
 - Pools Class B are pools with a TVL of less than \$50MUSD but more than 10MUSD
 - Pools Class C are pools with a TVL of less than \$10MUSD
4. Display the pools in a table with the different categories and sort them in descending order with the highest APYs first.
5. Create a simulation which maximizes the Projected Weighted APY using an allocation strategy with the following rules:
 - Pools class A should have a minimum allocation of 35%
 - Pools class B should have a maximum allocation of 35%
 - Pools class C should have a maximum allocation of 30%For this simulation suppose that the team is investing \$100,000 USD in 3 pools, the simulation output should tell the team in which pools is best to invest based on the above rules and the weighted combined APY of all of the pools.

Suggested Tools:

- Python 3+
- Jupyter Notebooks
- Streamlit
- Example Repository (<https://github.com/mattia-bolognesi/doomys>), this is an example repository that could be used from a third party but it does not have the data needed, it is only mentioned as a reference.
- Source of Data (<https://coindex.com/>)