

# Project Requirements and Guidelines

## Background:

White Rock is a 300 year-old company formed in London and now in 32 markets globally, including 8 in Asia. They are primarily an asset manager, with institutional clients as well as intermediaries who distribute their investment products.

Given the number of markets it operates in, White Rock has invested in its ability to manage business change across its business units, such as changes to shared services or regulatory driven change in processes. White Rock values delivering a great experience to its clients and employees and so has developed digital, cognitive science and data science capabilities.

White Rock collects a lot of internal data from various operational transactions, processes as well as activities monitoring. However, there is a desire for automated and intelligent process or technology to further mine the value of such internal data to the organisation in Singapore or any market.

Currently, the process to answer key Operational, Investment Management or Distribution related questions still requires a lot of manual intervention - searching for data, sense making, and communicating insights or recommendations. This may involve one or more of investment managers, traders, data insights team and relationship managers (RM), which is time consuming and not necessarily accurate, or complete. This impacts the organisation's ability to understand if it is operating at its best capacity, effectively or efficiently, and where the possibilities of opportunities to improve are, on its journey to becoming the leading digital asset management firm globally.

## What We are Looking For:

A POC (based on publicly available data sources) that will show us how to make faster and more informed decisions, to improve our business across Operations, Distribution or/and Investment Management teams.

### For Operations:

- Better resource planning.
- Identifying which processes are sub optimal.
- Enabling us to prioritize which processes to automate and/or digitize.
- Identifying opportunities for innovation.

### For Distribution (Sales and Marketing):

- Understanding cost to acquire, serve and retain our clients.
- Identifying indicators of customer churn and actions we could take to reduce it.
- Predicting future buying or selling patterns.
- Calculating optimal timing/ mechanisms for client contact.

### For Investment Management:

- Improving regulatory compliance processes.
- Enhancing identified processes for better efficiencies.

Below are some use cases that illustrate the primary uses for the prototype solution to address:

- When on boarding a new client, the solution speeds up due diligence process and quickly integrate into internal systems and processes.
- Where we believe the process to execute a portfolio is too long – look at data insights to pinpoint opportunities to improve, risk areas, activity costs.
- In a regulated environment, insights to know how we can respond to regulatory change – the time taken, resource usage, key activities, costs.
- Our Head of Distribution asks, 'what is the cost of client servicing'?
- We want to know the top 5 'broken' processes to fix.
- We want to identify the 'moments of impact' across our client interactions that drive success metrics, such as retention, spend.

Your team will choose an area [perhaps one bullet point above] and then propose and present your solution POC to White Rock Senior Management that are mostly not analytics trained and one data scientist. Know that many teams will be presenting and some may be presenting on the same area. What's so special about your solution that we should shortlist your solution to the next stage?

### On Datasets and Analytics\* Techniques:

Your solution may consists of several ideas or components. You will need to find at least one suitable dataset to demonstrate at least one idea or component. You are not expected to find perfect datasets or datasets that illustrate every idea.

As this is a graded project for this course, the main technique(s) used in your solution and POC must be learnt in this course.

### Data Sources:

You can use any of the following datasets or/and any other suitable data that your team found.

- IBM HR Data: <https://www.kaggle.com/dgokeeffe/ibm-hr-wmore-rows>
- HR Data: <https://www.kaggle.com/rhuebner/human-resources-data-set/data>
- Bank Churn: <https://www.kaggle.com/shrutimechlearn/churn-modelling>
- Personal Loan Classification: <https://www.kaggle.com/itsmesunil/bank-loan-modelling>
- Loan Eligibility Prediction: <https://www.kaggle.com/ninzaami/loan-predication>
- Mutual Funds: <https://www.kaggle.com/sanketc08/mutual-funds-eda>
- Netflix Stock Price: <https://www.kaggle.com/jainshukal/netflix-stock-price>
- Investment Survey: [https://www.kaggle.com/nitindatta/finance-data?select=Original\\_data.csv](https://www.kaggle.com/nitindatta/finance-data?select=Original_data.csv)

## Project Submission Deliverables:

1. A concise report in word document format, with a one page executive summary of your most important findings. Marks will be deducted for unnecessary paragraphs that waste the examiners' time to read. Include your team number and team members on the front page. Guideline: Max 20 pages, excluding front page, appendices, and executive summary.
2. A deck of PPTX presentation slides that summarize your important idea(s) and POC results. State the Problem/Opportunity statement. Include your team number and team members in the first slide, and **display the name of the team member who will be presenting the slide, at a corner in each slide**. This will be used for Individual Presentation Assessment in Project. Guideline: Max 25 minutes presentation in total.
3. Provide link to publicly available data and consolidate the final, cleaned dataset used, preferably in CSV format. Document the meaning of all the variables and coded data values in a separate data dictionary document.
4. The R script used to execute the analytics.
5. Presentations will be done in person in week 12 or 13 class. All team members must do presentation.
6. Put 1 to 4 in a zipped folder (without password) and upload your single zipped file into NTULearn class site. Include your team number in the zipped folder filename. If your dataset(s) or videos exceed data size limits, include links to download the datasets or/and watch the videos.

Note: Marks may be deducted for non-compliance to instructions or late submission.

## Project Grading Criteria and Submission Deadline:

Refer to Project Assessment Rubrics in the Assessment Components and Rubrics PDF.

Submission Deadline: End of Week 11 (Sunday).

Submission Online: NTULearn Class Site.

As there could be many teams (depending on enrollment) and we need to reserve additional 10 mins for Q&A for each team, some teams may be randomly scheduled to do presentation in week 12 class while other teams will do presentation in week 13 class. You and your teammates will only attend your scheduled timeslot. To be fair, all teams must submit all the deliverables by the submission deadline.

You are allowed to submit a revised corrected version of your slides or/and code or/and report after presentation and before end of week 12. Inform your instructor by email and do not delete the previous version in NTULearn as the timestamp serves as official record of your submission before the deadline.