

INTERNSHIP: INTERIM PROJECT REPORT 1

Dear Intern

Interim project report is an inherent component of your internship. We are enclosing a reference table of content for the interim project report.

The key objective of this report is for you to capture how far you have got in completing the internship work against milestones expected to be achieved within a specific duration and seek the mentor's feedback. Depending on the internship project and your progress (IT/Non-IT, Technical/Business Domain), you may choose to include or exclude or rename sections or leave some sections blank from the table of content mentioned below. You can also add additional sections. You can refer the project presentation to view the milestones related to your internship project. Please populate milestone# (1 / 2 / 3) and the milestone description in the interim project report based on the milestone for which you are submitting the interim project report.

You can refer the project presentation to view the milestones related to your internship project.

Internship Project Title	<i>Forecasting System - Project Demand of Products at a Retail Outlet Based on Historical Data.</i>
Name of the Company	<i>ICT Academy of Kerala</i>
Name of the Industry Mentor	<i>Debashis Roy</i>
Name of the Institute	<i>ICT Academy of Kerala</i>

Start Date	End Date	Total Effort (hrs.)	Project Environment	Tools used
<i>20 Apr 2023</i>	<i>15 Jul 2023</i>	<i>23.5</i>	<i>VS Code, Jupyter Notebook.</i>	<i>Excel, Python-3.9.13, Python packages like NumPy, Pandas, Seaborn, Matplotlib.</i>
Milestone	01	Milestone:	Student should be able to create a dataset. Clean the dataset and sanitize it.	

TABLE OF CONTENTS

	TITLE	PAGE NO.
(i)	<i>ACKNOWLEDGEMENT.</i>	3
(ii)	<i>OBJECTIVE.</i>	4
(iii)	<i>INTRODUCTION / DESCRIPTION OF INTERNSHIP.</i>	5
(iv)	<i>INTERNSHIP ACTIVITIES.</i>	6
(v)	<i>CHALLENGES.</i>	7
(vi)	<i>REFERENCES.</i>	8

ACKNOWLEDGEMENT

I *Govind S* would like to express my sincere gratitude to my Industry mentor *Debashis Roy* and *TCS ioN* for providing me with all the necessary resources and guidelines for the completion of this project Milestone 01 as part of the RIO-125 internship on Demand Forecasting.

I would also like to thank my colleagues who have helped me in various ways, such as sharing their knowledge and perspective, providing technical supports, and on giving me constant feedback of my works.

Finally, I would like to express my appreciation to the kaggle user *Pavan Kumar D* who has provided this dataset.

Thank you all for your support and encouragements.

OBJECTIVE

Create a forecast model by applying concepts of time series forecasting, quantitative forecasting methods, naive approach, ARIMA, etc.

The project milestone 01 involves several steps like: loading the dataset to the python environment in an IDE, cleaning and preprocessing the data variables and understanding data properly through univariate and bivariate analysis.

INTRODUCTION

In today's market of fast-paced service industry, predicting client/ customer demand for products/ requirements is a crucial for inventory management and meeting customer needs.

In this project, I will be developing a forecasting model that can predict the demand of products at a retail store based on their past data of timeline between 2012-2014.

Once the data understanding is completed and preprocessed, the next common procedure would be to select a suitable/ ideal forecasting model for the. respective After selecting the model, I will fit the model with the dataset and make predictions. To validate the accuracy of the predictions, I will visualize them using various visualization techniques.

Source of data:

The dataset was downloaded from the open source Kaggle platform.

About Dataset: Predicting the Demand of Products Across Stores of a Retail Chain.

A large Indian retail chain has stores across 3 states in India: Maharashtra, Telangana, and Kerala. These stores stock products across various categories such as FMCG (fast moving consumer goods), eatables / perishables and others. Managing the inventory is crucial for the revenue stream of the retail chain. Meeting the demand is important to not lose potential revenue, while at the same time stocking excessive products could lead to losses.

In this problem you are tasked with building a machine learning model to predict the sales of products across stores for one month. These models can then be used to power the recommendations for the inventory management software at these stores.

The datasets are provided as cited below.

train_data.csv Features:

- date: The date for which the observation was recorded.
- Product-identifier: The id for a product.
- Department-identifier: The id for a specific department in a store.
- Category-of-product: The category to which a product belongs.
- outlet: The id for a store.
- state: The name of the state.
- sales: The number of sales for the product.

Auxiliary Datasets:

- product-prices.csv: The prices of products at each store for each week.
- date-to-week-id-map.csv: The mapping from a date to the week_id.
- sample-submission.csv: The format for submissions.

INTERNSHIP ACTIVITIES

- ✓ Gone through the Welcome Kit and understood the Day Wise Plan.
- ✓ Passed the cut-off mark of RIO Pre-Assessment on my first attempt.
- ✓ Introduced myself in the DDR among my peer group.
- ✓ Watched and learned the youtube videos endorsed about forecasting methods, including 'Forecasting Methods', 'Moving Averages', 'Time Series Forecasting', and other models also that are part of the project reference materials.
- ✓ Data Mining/ Extraction through various open data sources. Finally found one and downloaded the dataset from Kaggle, loaded in excel for prior visualization before proceeding with the tools.
- ✓ Working on the dataset with my VS code IDE.
- ✓ Imported the dataset using necessary python libraries.
- ✓ Ensured that there are no missing values in the dataset.
- ✓ Converted the data types into the required formats.

CHALLENGES

One of the main challenge involved up till this now was on finding the ideal dataset that meets the TCS iON's project-dataset guidelines.

Once the dataset was obtained, the next challenge was on analyzing and to gain the proper insights/ understanding of data on which the various forecasting methods to be employed and in identifying the feature selections for creating forecasting models as I am not well versed with the domain. I would not say building the necessary domain knowledge is a hectic task, but it was a one of the small challenges that I faced.

REFERENCES

Link to the Dataset/ Data Card:	Predicting the Sales of Products of a Retail Chain
Link to Code/ Executable File:	Work file.ipynb