engagement proposal

Store Analytics

DEPARTMENT : Data Science

PROJECT NAME : Datascience Engagement 1.0

VERSION : 0.1

DOCUMENT RELEASE DATE : 27st - April - 2016

DOCUMENT VERSION : 1.0

DOCUMENT RATING : **Confidential**

Contents

[1 Client Background 4](#_Toc445305721)

[2 Client objective 4](#_Toc445305722)

[3 Proposed Solution 4](#_Toc445305723)

[3.1 Modeling Framework 4](#_Toc445305724)

[3.2 Modeling Process 5](#_Toc445305725)

[3.3 Model Hypothesis 6](#_Toc445305726)

[4 Duration of Engagement 14](#_Toc445305728)

[5 Engagement Model 14](#_Toc445305729)

[6 Roles and Responsibilities 14](#_Toc445305730)

[7 Model Validation And Refinement 14](#_Toc445305731)

[8 Key Phases 14](#_Toc445305732)

[9 DELIVERABLES 15](#_Toc445305733)

[10 Tools & Technology 15](#_Toc445305734)

[11 Tools & Licenses 16](#_Toc445305735)

[12 Support 16](#_Toc445305736)

[13 Entry Criteria 16](#_Toc445305738)

[14 Valiance Solutions 16](#_Toc445305739)

[14.1 Overview 16](#_Toc445305740)

[14.2 Valiance as analytics partner 17](#_Toc445305741)

[14.3 Leadership Team 17](#_Toc445305742)

[14.4 Awards 18](#_Toc445305743)

[14.5 Clients 18](#_Toc445305744)

[14.5.1 List of clients presently serving 18](#_Toc445305745)

[14.5.2 List of BFSI companies served in last 3 years 19](#_Toc445305746)

[14.6 References 19](#_Toc445305747)

Confidentiality Statement:

Copyright © ,Valiance Solutions Pvt Ltd. All rights reserved. This product or document may not, in whole or in part, be copied, photocopied, reproduced, translated, or reduced to any electronic medium or machine readable form, by any means electronic, mechanical, photographic, optic recording or otherwise without prior consent, in writing, of the copyright owner.

# Document Change History

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Version | Reviewed By | Modified By | Description of Changes | Date |
| 0.1 | Shailendra Kathait | Siddhant Verma | Draft Proposal | 27st April 2016 |
| 0.2 |  |  |  |  |
| 0.3 |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

# 1 Client BAckground

Store Analytics .

# Client objective

Store Analytics wants to forecast the sales for each of their store. The store sales vary based on the temperature, rainfall and other factors. The client wants a system that takes in input the parameters that affect the sales as well as the number of days for which the forecast has to be made.

Also, the system has to be deployed as user interface API for ease of usage.

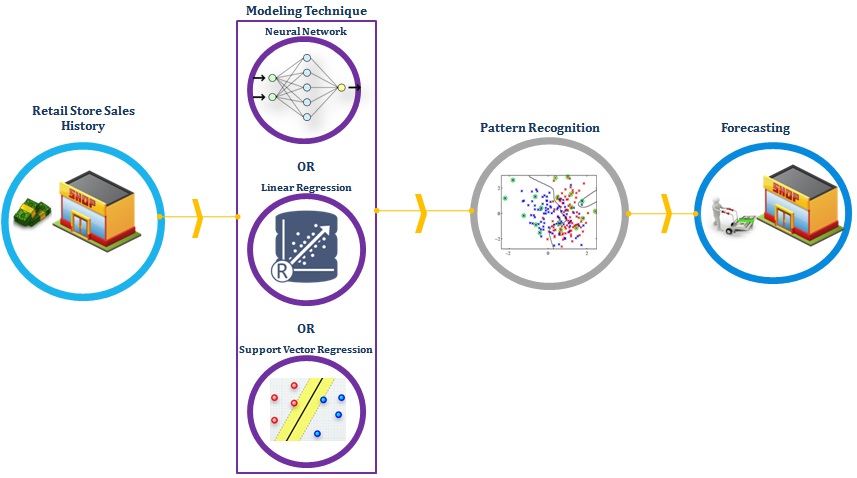
# proposed Solution

Valiance proposes a Quantitative framework development based on Machine Learning and front end development based on ASP .net APIs. Overall solution will comprise of following components

* ML based forecasting algorithms which would provide necessary predictions.
* ML based forecasting algorithms would consider the user input on to make predictions based on interaction of different chosen parameters as well as their values.
* As an output, ASP .net API would churn out user level interaction front end and the back end would combine with the ML algorithms to generate a output forecasting table.

**We will use internal data sources comprising historical retail store sales details and other available data like, temperature, rainfall, oil prices and interest rate at time of analysis to make recommendations.**

## Modeling Framework



## Modeling Process

Building & Validating the model

1. We will treat each store as separate segment to increase forecasting accuracy.
2. Separate Forecasting Algorithm will be developed for each store.
3. Dataset will be divided into Model building and Validation sample.
4. Exploratory data Analysis followed by Outlier treatment, Missing Value Treatment
5. Sample Independent Variables used for predicting sales are as below.
   1. Sales Lags for 1-30 days
   2. MA Sales Variables
   3. Create lag variables for Gas Price, Rainfall and Temperature
   4. Transformation of variables to improve fitment.
   5. Interaction Variable Creation
6. Daily sales will be used as dependent variable
7. Variety of supervised learning techniques will be used based on Exploratory data Analysis (Linear Regression, Neural Networks, SVR etc)
8. Model will be finalized based on MAPE on both Model and Validation sample.
9. Build the model on existing dataset. 70% of the dataset will be used to build the model
10. Validate the model against 30% of the dataset. Rebuild if performance characteristics don’t match.
11. Model once validated will be used to optimize sale forecast results.

Used for pre screening

Store Historical Sales

Data

* Retail store sales history
* Other Information as relevant (e.g., holiday details, day of week when store is closed)

Preliminary Data Analysis

* Check for missing or wrong values in variables
* Variable reduction & Transformation (correlation b/w continuous, frequency graphs, multi -co linearity checks)
* Data preparation for forecasting.

70% data for analysis

* Store’s transaction history
* Predict days with maximum sales.

Forecasting

Model

Validate the model against 30% dataset

Validated?

No

Yes

Rebuild

## Model Hypothesis

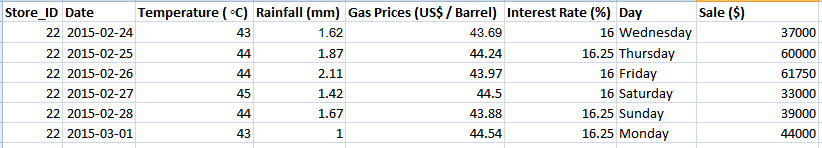
We propose to work on the model into 2 stages,

1. **Analytics**

In this part we intend to run Neural Network, Linear Regression and Support Vector Regression algorithms on the historical data of each retail store.

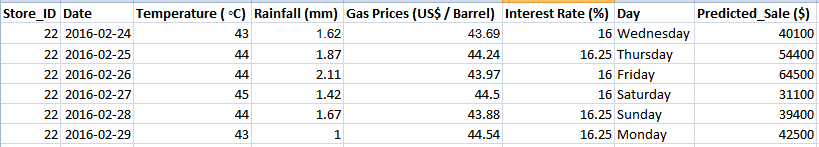
Consider for example the store in New York with Store\_id 22 has following historical data for Feburary,2015.

**Historical Sales Data :-**



Based on one of the forecasting techniques, we considered all the factors ( temperature, interest rate, rainfall and gas prices) to predict the sales for Feburary,2016. The results are,

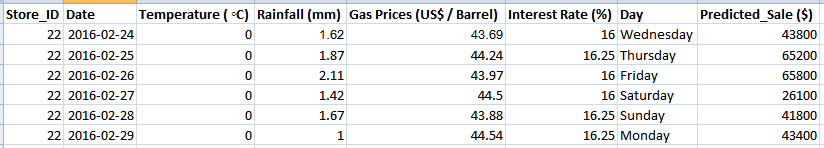
**Predicted Sales:-**



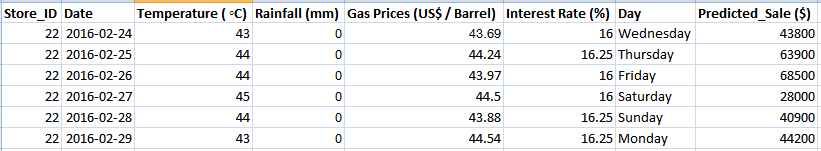
**Simulation**

We can make simulations over the model as well, by choosing a set of factors. The Predictions would vary with different set of input factors. Example,

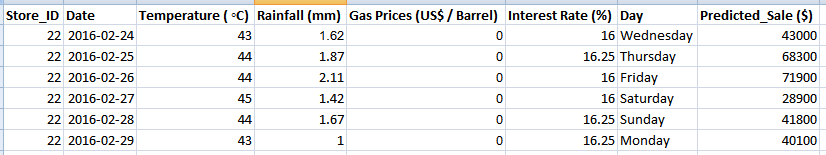
1. No Temperature



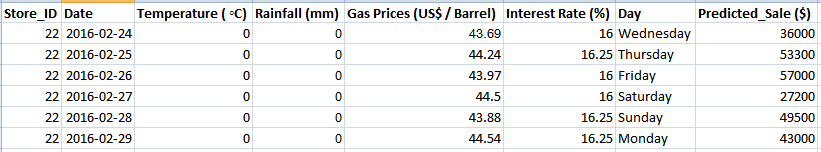
1. No Rainfall



1. No Gas Price



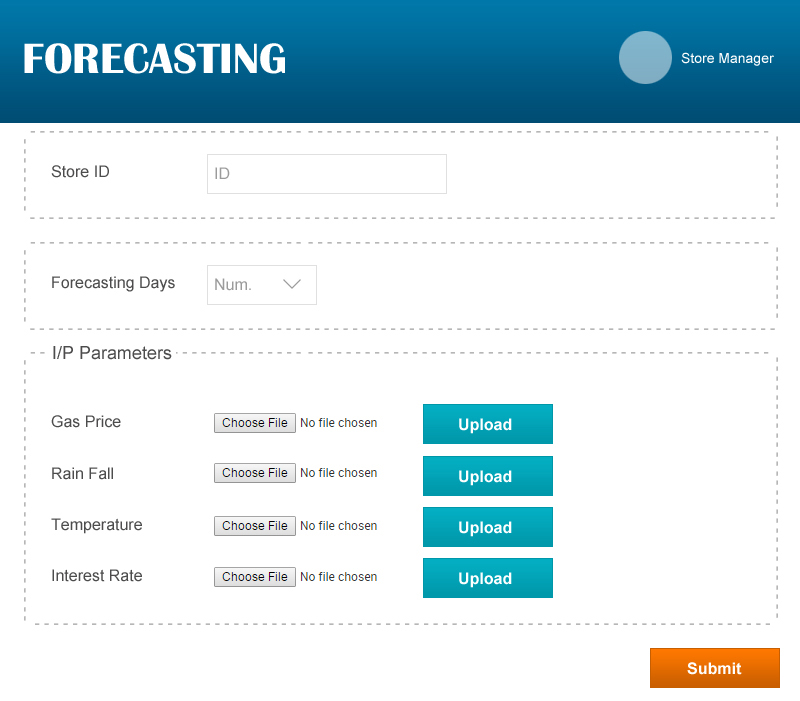
1. No Temperature and no rainfall



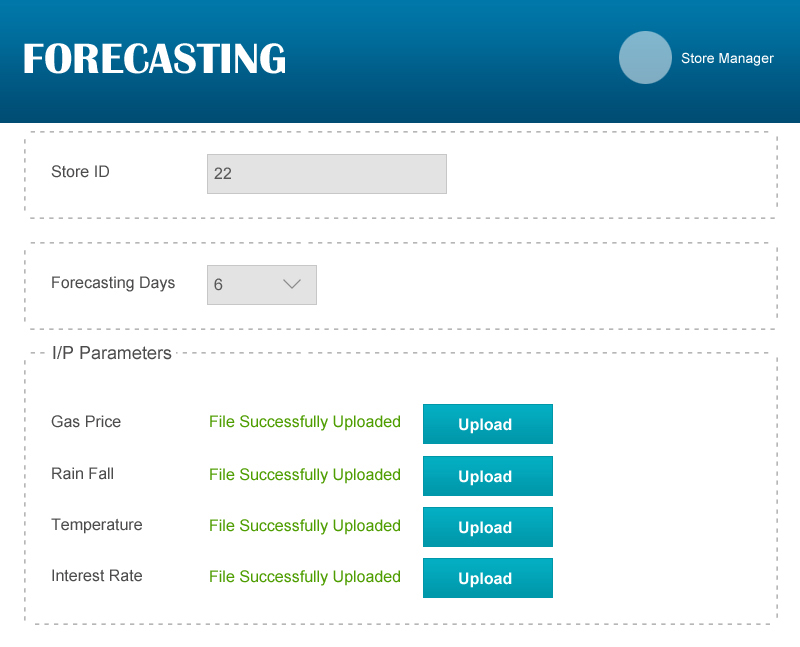
1. **Deployment**

This part involves using ASP.net API to design interface for the store managers. The interface or front-end of the system would allow the store manager to login and select his Store\_id. The system would allow him to choose the number of days for which he wants to predict the sales for as well as the factors/parameters he wants to consider for forecasting. A csv file has to be uploaded for each of the factors by the store manager.

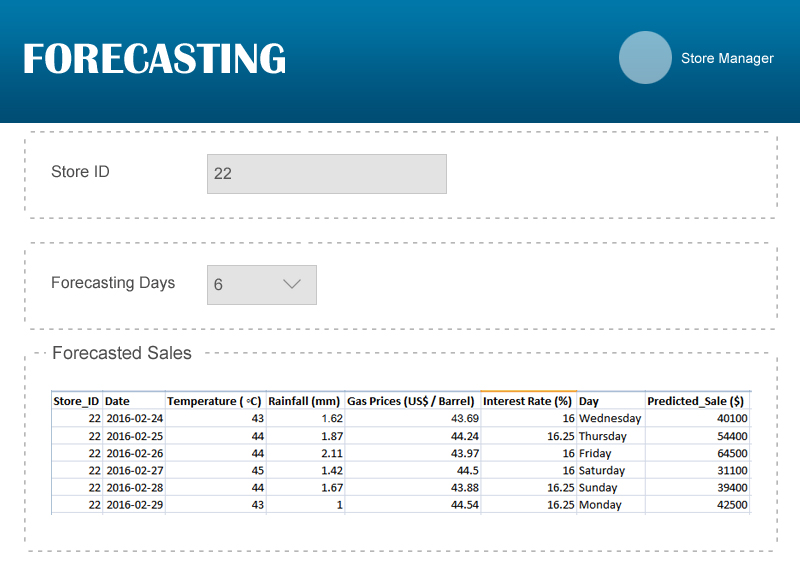
The home page of the API would look like,



Considering the example in the previous example, The store manager wants to predict the sales for the next 6 days for the store\_id 22, the screen would look like,



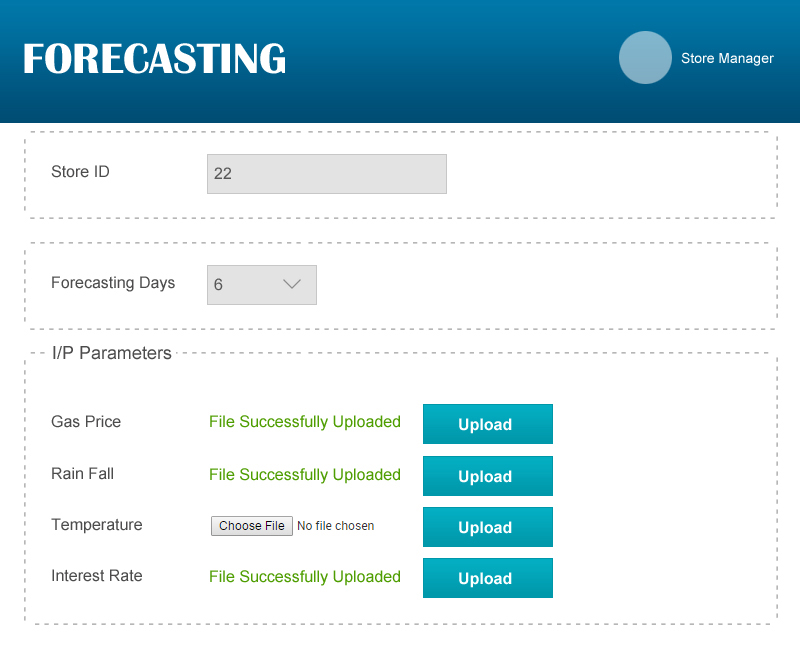
Now the result for input would look like,



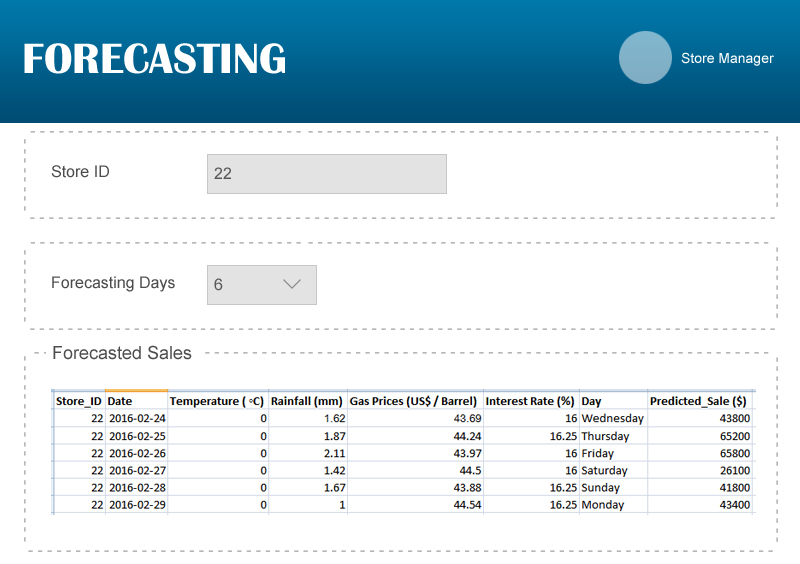
Again, if the user chooses to eliminate the factors, the input screen and output screen would look like,

1. No Temperature

Input :

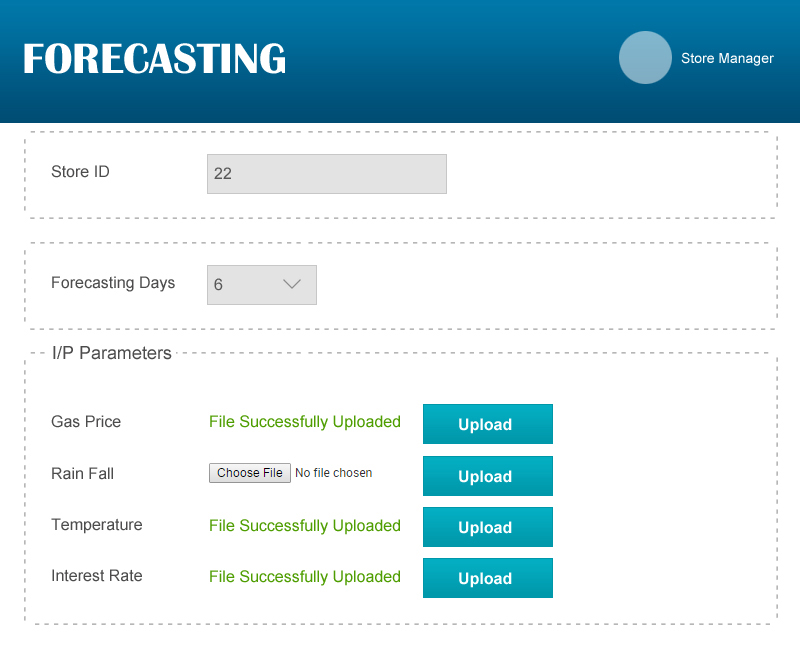


Output :

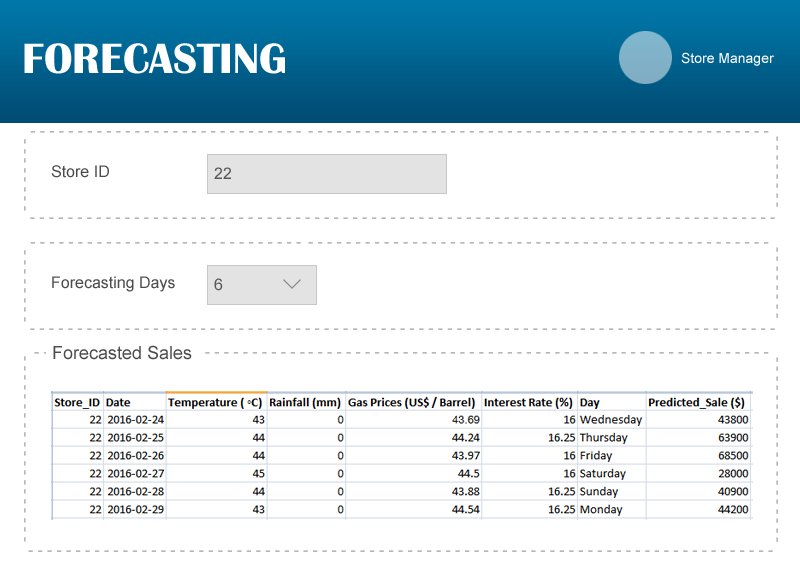


1. No Rainfall

Input :

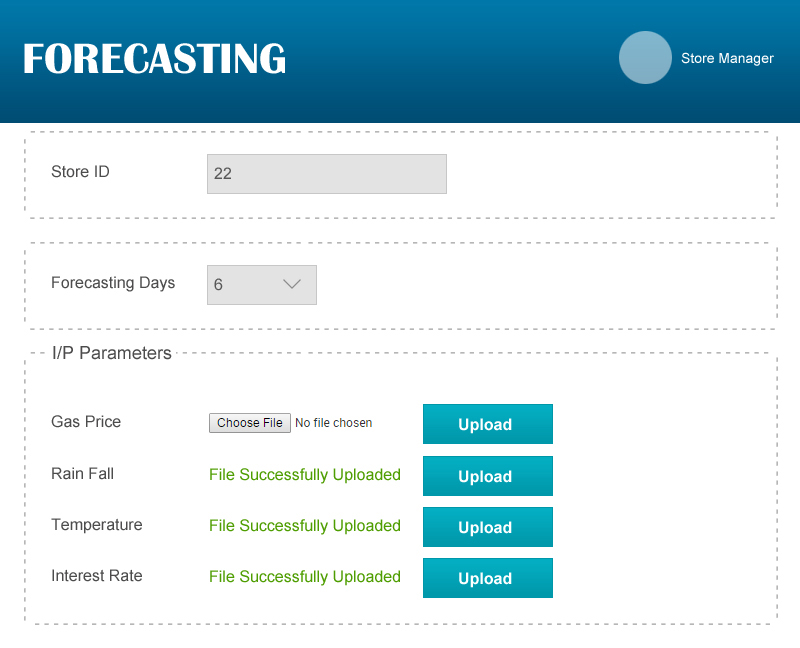


Output :

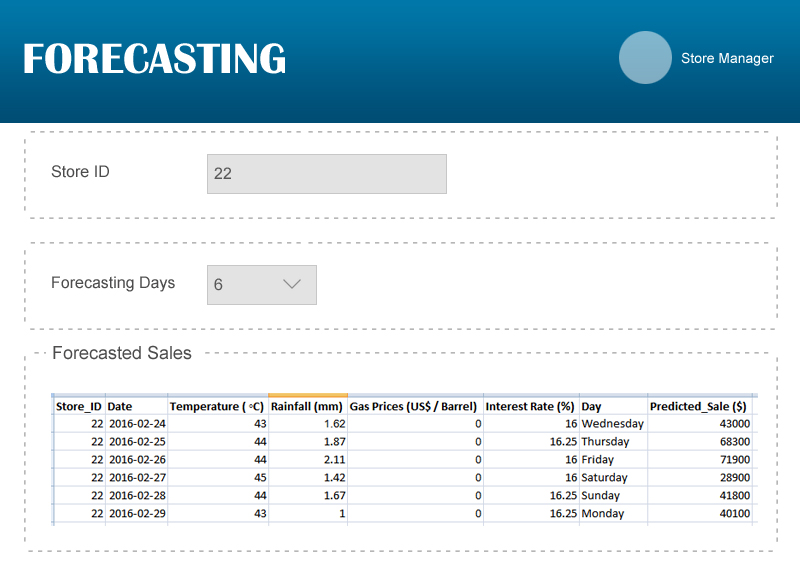


1. No Gas Price

Input :

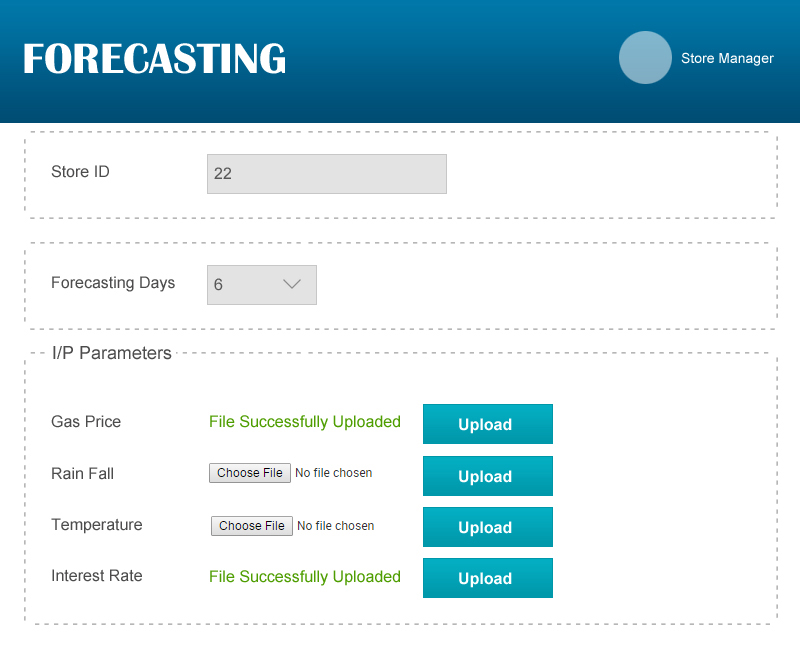


Output :

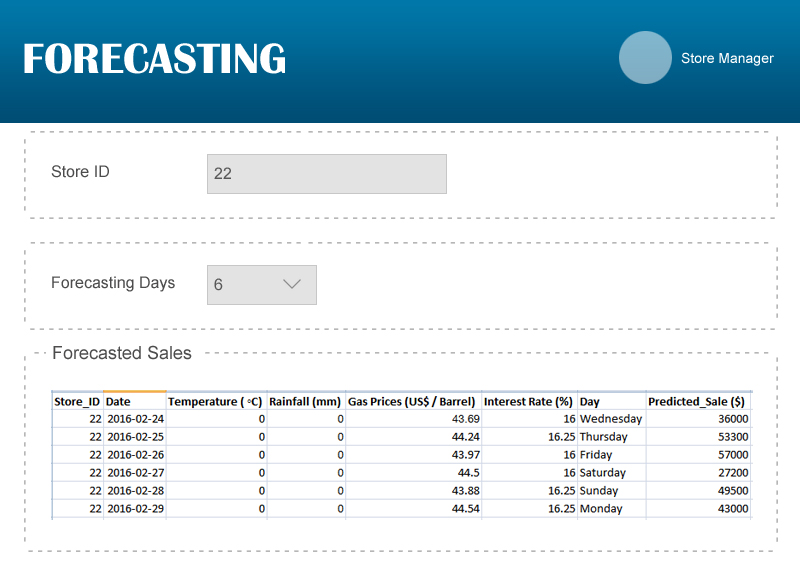


1. No Temperature and no rainfall

Input :



Output :



**Note :- The zero in the factor column occurs as the factor was not considered in forecasting.**

# DUration of Engagement

* Total duration of the project would be 6-8 weeks
* 6 months for model development.
* 2 months for model integration.

# Engagement Model

We can sort offshore, onsite or mix of both. Commercials will vary as per the model.

# ROles and Responsibilities

Valiance team will be responsible for following set of activities

1. Build model to predict sales.
2. Integration of the forecasting framework with the front end so as to provide each store manager to give input and provide him insights that he/she seeks.
3. Validate the efficacy of the framework over a period of time.
4. Review and iterate the framework if required.
5. Provide post implementation training & support on need basis.

Client’s Team will be responsible for following set of activities

1. Providing data for modeling. Data can be given either through access to database, or export to a csv or excel file.
2. Providing necessary IT and functional support to implementation and usage of algorithms.

# Model Validation And Refinement

Valiance will work with the client to improve accuracy of algorithms over the support period. This will include

1. Monitoring current algorithms for performance and tweaking them if required.
2. Develop additional algorithms if necessary.

# Key Phases

1. Data Discovery
   1. Understanding data availability within IT systems
   2. Preparing Data Dictionary
   3. Data Quality Assessment
2. Data Integration
   1. Aggregating retail store data and other relevant data (like temperature, rainfall, gas prices, etc) needed for analysis.
3. Initial Event Trends.
   1. Discover trends of various factors in relation to predict sales in forthcoming days. Involves exploring correlation through statistical measures, graphs.
4. Build Machine learning algorithms for predicting best offers/deals
   1. Different ML algorithms from Linear Regression to Neural Networks to be tested.
5. Implementation
   1. Integration of algorithms with user interface and other IT systems as per business use case.
   2. Technical document of algorithms.
   3. Training of client’s team in usage and upkeep of algorithms.
6. Monitoring
   1. Evaluating performance of algorithms in making predictions and understanding user input on frontend.

**Timeline for each phase to be decided after discussion with the client.**

# DELIVERABLES

1. Final Report on the exercise covering following information
   * Significant trends & factors affecting the forecasting of sales made by each retail store, including details of input on the frontend.
   * Performance of the model against out of time data. Report will be further augmented with each month’s validation results of the model for continued period of time fixed after implementation and with discussion with client after completing the first phase.
   * Details on techniques/algorithms used to build the model.
   * Detailed instructions to business users on usage of the model output.
   * Instructions for the technical team on validation of the model performance and stability monitoring.
   * Score card based on business objective
2. Programmatic implementation of the model(s) ( in discussion with client’s IT team)
3. Training to client’s Business & Technical Users for usage and validation/monitoring of the model.
4. Implementation & testing support.

# Tools & Technology

|  |  |
| --- | --- |
| Operating System | Linux |
| Data Storage (if needed) | MySQL/ SQL SERVER |
| Data Science Platform | Python/R |
| API’s | Python/ ASP .net |

We are open to building and implementing algorithms using any proprietary software products from SAS, IBM or Microsoft.

# Tools & Licenses

1. Any proprietary licenses and tools required for the project will be purchased by Client and provided to Valiance.
2. Client will provide necessary infrastructure required to host the algorithms. Additionally hardware infrastructure for development needs to be provided by client in case of onsite engagement.

# Support

1. Valiance will provide continued support to the client post delivery. This will cover any iteration required in the model if the performance is not as per expectations and out of time validation. This also includes model validation and performance report for management.
2. Support will be provided through both offshore and onsite presence as and when required.
3. Model may not perform as expected if there is a significant change in current population characteristics and the one on which model was built. This is measured by PSI (Population stability Index). In such a case model recalibration will be required and will fall outside the scope of support.

# Entry Criteria

1. MSA is signed
2. NDA signed
3. SOW signed

# valiance solutions

## overview

Valiance Solutions is a data engineering and advanced analytics firm helping clients create new products and decision making capabilities using Big Data technologies and data science. We work with Marketing, Product Development and Operations leadership create strategic and tactical solutions that regularly deliver business growth.

Our diverse team of data scientists, data engineers and data visualization experts create outcomes that consistently outperform baseline performance through a rigorous test and learn approach in model development and with the implementation of machine learning solutions that ensure continual improvement.

Valiance has offices in New Delhi, India and New York City, U.S

## valiance as analytics partner

Valiance brings in analytics capabilities across entire data life cycle with data engineering and machine learning. Our success lies is building complete end to end analytics platform across domains that deliver predictive intelligence capabilities. This differentiates us strongly from other analytics firm which are either consulting focused or software development companies.

We have worked on several exciting projects at intersection of Big Data, Machine Learning and Natural Language Processing. Key projects executed in year 2015 include

* Machine Learning algorithms for Programmatic Media buying platform for Mobile Advertising Company. Algorithms predict propensity of customer to click on a display ad based on browsing behavior across partner websites.
* Data driven consumer lending platform for Housing Finance Company in India. Project is development phase since last 2 months with plan to roll out 1st version in Feb’16.
* Data platform for ingesting and processing 6TB+ plus unstructured data using text mining and NLP. This platform is used to run Machine Learning algorithms for predicting customer attributes (income, gender, location etc) based on transactional information.
* Identification of tax payers profile for potential tax evasion using demographics, employment, financial profile and tax filing details.
* Lead prioritization analysis for Education service provider in US. Project involved building predictive scoring engine for scoring applicants based on demographics, interaction with company’s various digital touch points.

## Leadership Team

***Vikas Kamra (Co-founder & CEO)***

Graduate IIT Delhi

Vikas heads business development and marketing for Valiance. Prior to founding Valiance he worked with investment banks like Merrill Lynch, Bank of America for nearly three years as a technology consultant. He also had a 1 year stint with public opinion firm setting up their data and analytics systems which led to birth of Valiance.

***Lewis Tierney (VP, Client Success)***

MS (Integrated Marketing Communications), Northwestern University

Lewis Tierney is a customer-centric marketing analytics professional with a track record of implementing innovative solutions that facilitate better decision-making and drive business growth. He has extensive experience partnering with marketing organizations to create data-driven customer segmentation in a variety of industries including financial services, retail, business information services and not-for-profit. Lewis has also applied his analytic background in the customer service realm serving as Director, Global Voice of Customer Analytics for American Express for nearly four years

***Shailendra Singh (Co-founder & Head of Analytics)***

Graduate M.E.R.I, Kolkata, IIT Delhi

Shailendra Heads Analytics delivery for Valiance where he is responsible for building Machine Learning Products and Analytics driven outcomes for our clients. He brings 7 plus years of core Machine learning & Analytics experience with Fortune 100 companies like IBM, American Express & ICICI Group across EMEA, US and Indian Subcontinent region. Shailendra has deep Interest in ANN, Biological Networks & Optimization and holds several Patents and is Anchor author of several publications on Machine Learning & Optimization.

***Ankit Goel (Co-founder & CTO)***

Graduate IIT Kharagpur

Ankit oversees development of technology products for our clients based on big data, cloud ecosystem. This includes building visual analytics platforms, integration of machine learning modules .During his 10 years plus experience Ankit has worked in various roles of Product head, technology lead, software developer since his professional journey with Investment banks and startups.

## awards

* *Valiance Solutions has been named among the “Top 10 Emerging Analytics startups in 2015” by Analytics India Magazine.*
* *The winner of “Big Data & Analytics Awards” in the category “Most admired Big Data project”, year 2016.*

## Clients

### List of clients presently serving

1. Govt. of India (Income Tax Deptt)
2. Saral Dhan Home Finance (Housing Finance)
3. Prominent Indonesian Bank
4. Prominent Thailand Bank
5. Crif High Mark
6. Applied Mobile Labs
7. Digital Solutions Firm (Confidential)

### List of BFSI companies served in last 3 years

1. Saral Dhan Home Finance
2. Highmark
3. Aegon Life Insurance
4. Future Generali Life Insurance
5. Bajaj Finance

## references

To be shared on request