Spray-Tek Weather Project

Contents

[1. Project Introduction 2](#_Toc66045410)

[2. Version Control 3](#_Toc66045411)

[3. Contacts List 4](#_Toc66045412)

[4. Project Plan 5](#_Toc66045413)

[5. Meeting Summary 6](#_Toc66045414)

[6. Targets 14](#_Toc66045415)

[7. Data Visualization 15](#_Toc66045416)

[8. Current Problem 16](#_Toc66045417)

# Project Introduction

More specifically, given the dependence of their drying products and services on weather

conditions, the company is interested in using weather data to forecast the processing times for

different batches of SKUs.

The team is expected to:

• Develop a database with manufacturing data that will be used to develop statistical models

• Develop, manage and maintain statistical models, including but not limited to:

o Forecasting of processing times based on weather data

o Forecasting downtime and preventive maintenance issues

* Flavors run on dryers in NJ and Flavors run on dryers in PA, did they have different affect by weather
* Some production categories have different affect scale to humidity, we need to find which product is more sensitive.
* The dryer rate on different custom category are different. Find the best match custome product and Dryers.
* Our mission contain regression and classify part. For regression part, we need to predict Rate or Drying Time with mae and mse error. For classify part, we might need to use high dimension to classify costum product

# Version Control

|  |  |
| --- | --- |
| Version | Modification |
| V 0.1 | Initial document frame |
| V 0.2 | Add Feb 22 Meeting Summary |
| V 0.3 | Add Target section |
|  |  |
|  |  |
|  |  |
|  |  |

# Contacts List

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Sprary-Tek Weather Project Contact List | | | | |
| Name | Role | Email | Phone | Responsibility |
| Stevens Part | | | | |
| Alkiviadis Vazacopoulos | Professor | avazacop@stevens.edu |  |  |
| Pranay Bhandare | Student | pbhanda3@stevens.edu |  | Coordinate between Stevens and Spary-Tek |
| Jungho Park | Student | jpark28@stevens.edu |  | Project Plan |
| Yi Yi | Student | yyi2@stevens.edu |  | Platform Establish |
| Zi Wang | Student | zwang210@stevens.edu | 347-403-3376 | Project Document |
| Hao Yan | Student | hyan12@stevens.edu |  |  |
| Sprary-Tek Part | | | | |
| Eric Koelle | IT Manager |  |  |  |



# Project Plan

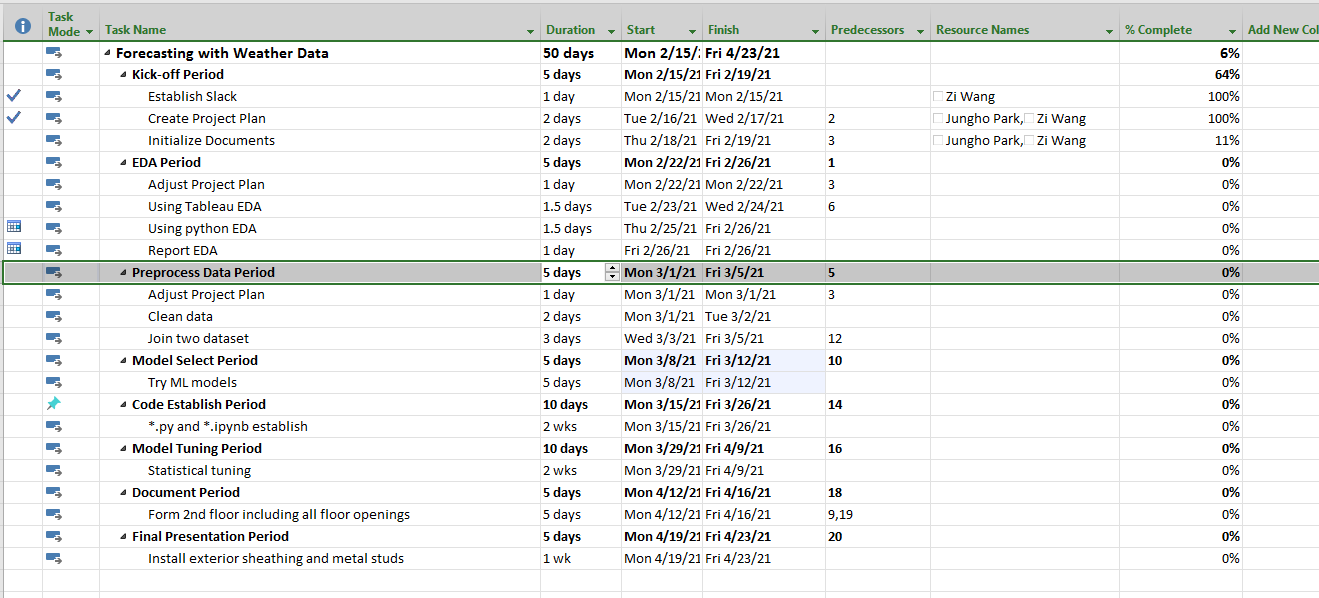
1. Stevens account can access VLE for Microsoft Project.
2. Different platform access Stevens VLE <https://sit.teamdynamix.com/TDClient/1865/Portal/KB/?CategoryID=11231>
3. Download corresponding Citrix version

https://sit.teamdynamix.com/TDClient/1865/Portal/KB/ArticleDet?ID=87051

1. Access Stevens VLE

<https://ra.vdi.stevens.edu/vpn/index.html>

1. Select Microsoft Project and Start using





# Meeting Summary

* 1. Feb 17 Meeting Summary
* Create slack
* Find Project Management tools (agile) - Project Planning by Friday
* Professor[Vazacopoulos](https://sit.instructure.com/courses/47171/users/11041) will put data on Slack / might ask for more data during progress
* Weather data & humidity -> using weather data to forecast the processing times for different batches of SKUs
* Visualization & descriptive analysis then do MLFind machine learning articles regarding such problems for implementation.The other 809 project is regarding production planning maybe we can connect both?
* Want to review / progress check-up every Monday
* Remind the recording to professor
* By Friday Meeting
* Zi: Slack setting
* Pranay: Connect with Spray Tek’s employee
* Yi Yi: Choose tech (e.g. python R Tableau)
* Jungho: Find Agile PM tools, possibly setup
  1. Feb 23 Meeting Summary
     1. Total eleven dryers
     2. [‘Moisture Targe’] is usually given to us by our customer and that is the moisture value in the dry powder once we have done processing it. For example, the customer set the target moisture value want to be 5%. This column is text field (string) that entered the requirement of customer by operators from Spray-Tek.

Request moisture target

* + 1. [‘Bulk Density’] is a measure of how much fit into a fixed volume of dryer, how much mass would be fit into a fixed body. This column is also a text field which is subject to the variability of the person entering the data in there. This is just for us to figure out how big box we can put it in. Then want to act 50 pound boxes, we need to make sure that it’s going to fit in.
    2. Dryer process, raw materials are some sort of starch or government material like gum acacia (what’s this?). There are typically mixed with flavorings which we may or may not have all the ingredient information because some are proprietary. We don’t make any product we just process for our customers. Typically, you have a starch powder, then you have sort of flavoring powder. We may use a flowage (Anticaking agent), like silicon dioxide, to help material flow better. If you need more detail, Eric can ask process engineer to try to get more info.
    3. [‘Flow’] There’s nothing empirical to this generation. Someone just say this flow is well and other flow is poorly. It’s just a matter of communication to people running on our equipment, how to expect this material to come out of the dryer. Operator have to pay attention to build up on the wall inside the dryer because it doesn’t flow very well. If you got build up on the chamber, you got hotspot and cause fires. Somebody with experience been here for 15 years have pretty much in charge of process in the whole time. He has a good idea of what will absorb moisture more readily than other things. But there isn’t any sort of concrete number behind that.
    4. [**‘Hygroscopicity’]**
    5. Is there any other data we can use?

We can set another meeting and talk about what other data is available. Eric just provide data that he thought would be useful

* + 1. How would you want to make runtime prediction in respect to weather data?

The weather effect either the rate or the time that it takes to run any specific product. There might be none affect on some product and might be a great affect on other product.

Dryer 04 is pretty much a fragrance only dryer so we don’t really have a way to tell it on data.

* + 1. Why there are so many outliers in our data?

You have to compare it to scheduled or the schedule dry quantity or the actual dry quantity. If A run is 20000 pounds, it’s always run much longer than running 5000 pounds. So, they can influence drying time. Dryers can run 24 hours.

You are looking at multiple different products. We are not running the same product on a dryer all the time, it’s different product all the time. You will see varying times and varying batch sizes.

* + 1. The two main things we look at here are rate and yield. Rate is better. Yield is how much losses we have as far as like build up in chamber. When you don’t run the dryer the right parameters and maybe you have too much humidity in the dryer.

Water in the badge, a lot of things can affect that.

* + 1. [‘CustItem’] Different material will run at different rate in the dryer. Anytime you have a different customer time, that is different material. The custom time number is the unique material code that we assigned to the products. For instance, on the first row, we can see [‘CustItem’] =10013-000 running on dryer 04. If that exists again in the table, they should have a very similar rate. Eric can give addition data about the category of customer item. Something like fragrance flavored, fragrance flavored is nutraceutical nutritional we have a couple of other categories. There are six or seven categories that we typically use. CustItem is just sequential number, we got a new product from customer we provide next number incrementally
    2. [‘BatchNumber’] the only difference is NJ and PA. Dryer01-04 in NJ, other in PA.
    3. Does each dryer has any specific purpose or just random assign?

Technically speaking, any dryer can run any product for the most part, however, in PA, we don’t do any fragrances. That’s just to keep any sort of orders of a flavor products. You don’t want you baby formula smell like Cologne.

Dryer09 runs very slow product.

Ask opertiona

* + 1. The whole process is removing moisture from a liquid to produce a powder. If the air coming into the dryer has more moisture in it. It’s gonna affect your rate. It shouldn’t typically affect your yield but it’s probably going to affect your rate.
    2. Is there any ventilation system inside the factory or dryer?

They do not use condition are. The air directly from outside. We don’t have any sort of de-humiditiy on the intal or dryer.

* + 1. What is the beginning percentage of moisture?

I could give you the solid percentage of the batch. Everything else is water

* + 1. Cleaning time not affect by the weather. We are more concern about dry time not total time. Total time can include downtime and downtime can be any reason like mechanical, process related. Downtime are not related to the performance of the product of the dryer.
    2. When you see yield percentage great than 100%, we had a blend back. When we run the dryers, the customer may give us a particle size expect. When some particles are larger than customer require, we use a shifter with a particular size screen through size mesh. Only the powder that meet customer’s requirement will fall through the screen and go into the box or the container. Everything else stay on top of the screen will be set into a separate container, which will be considered oversized material. If we are running multiple batches back to back of the oversized material, we dump it back into the batch and re-dry it again. We actually put in material that’s not accounted for in the batch and that’s how you end up with a yield over 100%. No more than 110% should be consider value data.
    3. Eric will provide custom item category, the flavors fragrance and nutritional. We can look at the categories of products and how they may be affected by weather
    4. The weather will directly affect [‘rate’], then [‘drying time’], last is [‘yield percentage’]
  1. Feb 26th 9 AM Outline

Good Morning everyone. Here is today meeting outline. We need to address topic together.

1. Next week plan.
   1. We have complete primary stage EDA, data clean and data merge for production sheet. We also need to do the same thing to our weather data.
   2. After a brief explore the weather data, I realized that this is a basic weather data. If we use this basic information to find correlation, it might not be significant. So, I offer a proposal that we might need find more professional weather data when we face analysis dilemma in the future.
   3. I believe there are someone who have tried to find weather influence on the faculty machine, not specific dryer but another helpful experience. We might need some time to search article, papers and website that help us start not from zero.
   4. After completing one of those tasks, we start our feather engineer, including converting description to numerical and numerical data aggregation.
   5. If Eric cam bring some operation engineers to our meeting, we can bring our questions to address more dryer one-hand information and experience.
2. Adjust project plan.

Every week we need to talk about our project progress and adjust project plan weekly. After completely new project plan, we will send update very Monday to our customer.

1. Fix weekly meeting time

For now, Tuesday 9 AM might be a good time for our customer meeting. We also need schedule a Stevens meeting before customer meeting.

1. Balance project and your study

I understand everyone have its own study pressure, working pressure and job-hunting pressure. If you have another interview, homework and work task which are more important than our project, feel free to do you things first. This is only a project, try to find your balance by yourselves. I believe our team members will support you.

1. Sometime each machine will run more than one day, so we need to find a better way to merge our data, for instance, increment every production record by hours. If single dryer running over 1 hours, we need increment this single line into multi-lines separated by hours. And

March 2nd Meeting Agenda

* + 1. Last work report
* Kick-off meeting
* Production sheet EDA
* Initial documents
  + 1. This work plan
* Weather data EDA Junho
* Advanced weather data research Yi Yi and Hao
* Clean and join Data Pranay and Zi
  + 1. Project Plan
* Communication List
* Project Plan weekly update
  + 1. Visit request

Because we not familiar dry technology, we want to physicality visit your NJ factory, communicate with your operation engineer, and know more about the potential factor that might affect dryer rate

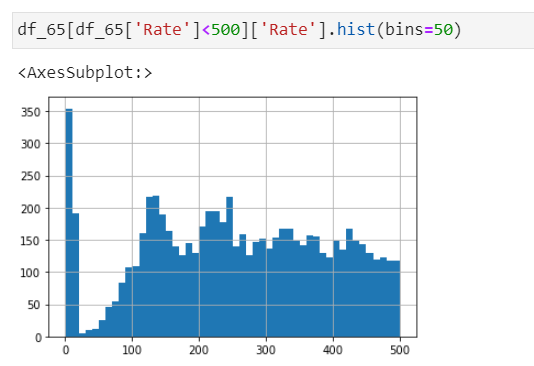
* 1. March 9 Meeting Agenda
     1. Last Week Report
* Yi Yi will report her weather report result
* Zi will report code progress
  + 1. This Week Plan
* Pranay will build software flowchart
* Hao and Zi, Yi will start primary data analysis
* Jungho will combine weather analysis result with data analysis
  + 1. Current Problem
  1. Dasf
  2. Sdf
  3. Asdf
  4. Asd
  5. f
  6. Coding
  7. Code Logically

Import\_data()->clean\_data()

# Data Visualization

# Current Problem

* 1. We find one Dryer 06 record in NJ (index = 5036), which should be in PA due to time(2019) and prodline(flavor)
  2. We find one Dryer 09 record in NJ (index = 4714) ), which should be in PA due to time(2019) and prodline(flavor)
  3. We find two Dryer 01 record in PA (index = 925, 1092)
  4. We find 23 Dry 11 records in PA, we intend not use them
  5. Dryer04 88.5% is Fragrance
  6. Draw software flow chart
  7. There are some [‘rate extremely high
  8. [‘Food Addit’] and [‘Food Additive’] same?
  9. PA are not fully processing [‘Flavors’]
  10. Rate = ‘Qty’ / drying time\_hrs. So we can only focus on [‘Rate’]
  11. From plot [‘Rate’] smaller than 20 is abnormal and this only occur in range(0,20) very strange



**Anomaly detection machine learning algorithm**