

## Instructions

### Instructions

- This homework assignment is worth 84 points.
- Please submit **.tfl** and **.twb** file for each exercise to Blackboard.
- **Please strive for clarity and organization.**
- **Due Date: April 8, 2024 by 11:59 pm.**

## Exercise 1

We are given information about the maximum daily temperature from a weather station in Flagstaff, AZ. This file (`weather_station.csv`) is in a wide format, where each row represents a month and the columns 1, 2, ..., 31 represent the day of the month the observation was made.

- (3 points) Load the `weather_station.csv` datafile to tableau prep.
- (6 points) Drop the first column, and pivot the data. The final data has only four columns: `Year`, `Month`, `Day`, and `Max Temperature` (see the below table).

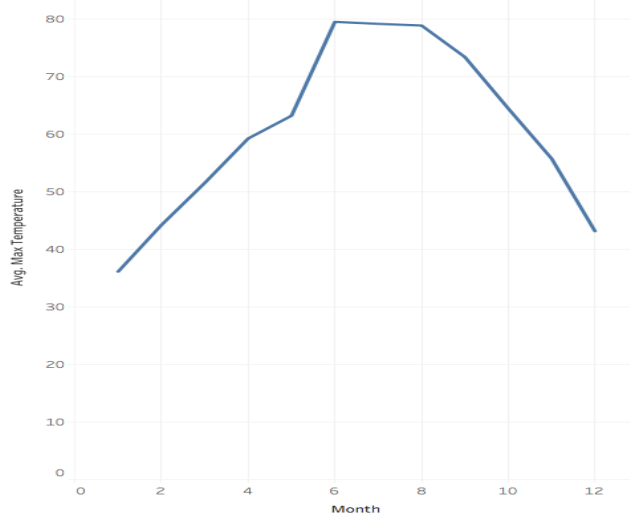
	A	B	C	D	E
1	Max Temper	Day	Year	Month	
2	71.06	1	1985	5	
3	62.96	1	1985	6	
4	80.96	1	1985	7	
5	77	1	1985	8	
6	82.94	1	1985	9	
7	64.04	1	1985	10	
8	50	1	1985	11	
9	33.08	1	1985	12	
10	42.98	1	1986	1	
11	48.02	1	1986	2	
12	62.06	1	1986	3	
13	62.06	1	1986	4	
14	66.02	1	1986	5	
15	75.02	1	1986	6	
16	75.92	1	1986	7	
17	87.08	1	1986	8	
18	71.96	1	1986	9	
19	60.08	1	1986	10	
20	55.04	1	1986	11	
21	48.02	1	1986	12	
22	48.92	1	1987	1	

Export this data to a `.csv` file called `weather_station_long.csv`.

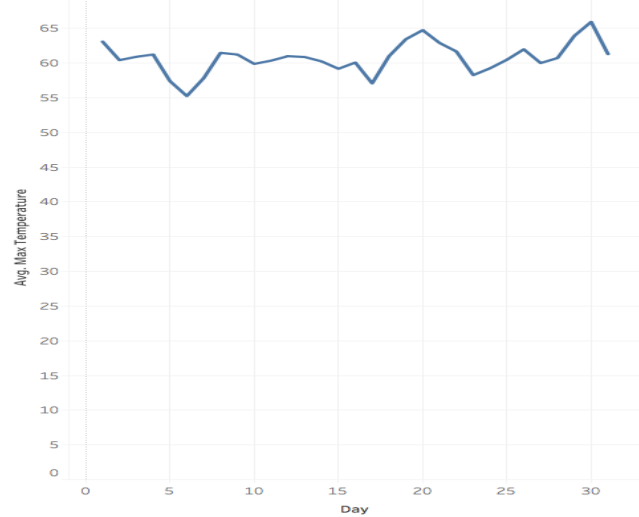
- (10 points) Using the `weather_station_long.csv`, duplicate the below dashboard.

Maximum Temperature 1985-2015

Average Maximum Monthly Temperature



Average Maximum Daily Temperature  
Year 2008



## Exercise 2

Consider the `refine.csv` datafile. Right away, you'll notice that the data set has a few issues:

- There are four brands: Philips, Akzo, Van Houten and Unilever. However, there are many different spellings and capitalizations of those names.
  - The product code and number are combined in one column, separated by a dash.
- (a) (3 points) Load the `refine.csv` datafile to tableau prep.
- (b) (5 points) Clean up the `company` column, so all of the misspellings of the brand names are standardized. For example, you can transform the values in the column to be: philips, akzo, van houten and unilever (all lowercase).
- (c) (5 points) Separate the product code and product number into separate columns i.e. add two new columns called `Product_Code` and `Product_Number`, containing the product code and number respectively.
- (d) (5 points) You learn that the product codes actually represent the following product categories:
- p = Smartphone
  - v = TV
  - x = Laptop
  - q = Tablet

In order to make the data more readable, add a column with the product category for each record.

- (e) (5 points) You'd like to view the customer information on a map. In order to do that, the addresses need to be in a form that can be easily geocoded. Create a new column **Full\_address** that concatenates the three address fields (address, city, country), separated by commas.
- (f) (3 points) Export the data to a .csv file called **refine\_clean.csv**.

## Exercise 3

Consider the **Transactions.xlsx** and **Targets.xlsx** data-files. The below picture shows a some of the rows of the daily transactions for each product. It contains the price, quantity and income.

	A	B	C	D	E	F	G
1	TransactionDate	ProductID	ProductName	Price	Quantity	Income	
2	3/1/17	1	Liquid Soap	1	76	76	
3	3/1/17	2	Bar Soap	1.5	26	39	
4	3/1/17	3	Candle	10	61	610	
5	3/1/17	4	Diffuser	15	59	885	
6	3/1/17	5	Hand Wash	5	98	490	
7	3/2/17	1	Liquid Soap	1	42	42	
8	3/2/17	2	Bar Soap	1.5	95	143	
9	3/2/17	3	Candle	10	30	300	
10	3/2/17	4	Diffuser	15	39	585	
11	3/2/17	5	Hand Wash	5	81	405	
12	3/3/17	1	Liquid Soap	1	15	15	
13	3/3/17	2	Bar Soap	1.5	6	9	

Also the finance team has provided a datafile that contains the sales targets. It is a weakly breakdown for 2020 by product.

	A	B	C	D	E	F
1	Year	Week	ProductName	Quantity Target	Income Target	
2	2020	1	Liquid Soap	158	158	
3	2020	2	Liquid Soap	473	473	
4	2020	3	Liquid Soap	473	473	
5	2020	4	Liquid Soap	473	473	
6	2020	5	Liquid Soap	473	473	
7	2020	6	Liquid Soap	525	525	
8	2020	7	Liquid Soap	563	563	
9	2020	8	Liquid Soap	563	563	
10	2020	9	Liquid Soap	563	563	
11	2020	10	Liquid Soap	563	563	
12	2020	11	Liquid Soap	570	570	
13	2020	12	Liquid Soap	570	570	
14	2020	13	Liquid Soap	570	570	

Using the **Transactions.xlsx** and the **Targets.xlsx** data-files, do the following:

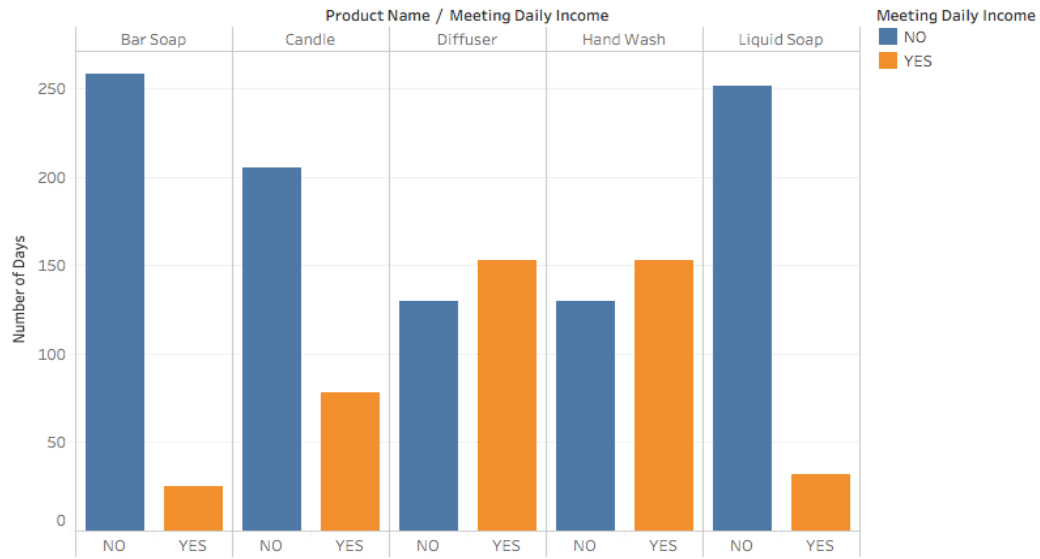
- (a) (3 points) Load both the `Transactions.xlsx` and the `Targets.xlsx` data-files to Tableau prep.
- (b) (8 points) Add a **Clean Step** to the `Transactions` table. Create two fields from the `TransactionDate` column: **Week** (week number) and **Year** (year number) (hint: use the `DATEPART` function). After, remove data from years 2017, 2018, and 2019; that is, only keep 2020 data.
- (c) (6 points) Add a **Join** step to the `Transactions` table. Conduct a left join using **Year**, **Week**, and **ProductName** as the key variable to conduct the join.
- (d) (10 points) Add a **Clean Step** to the resulting table from part (c). Remove redundant columns; that is, remove **Year-1**, **Week-1**, and **ProductName-1**. After that, create two new fields: **Quantity Daily Target** and **Income Daily Target**. Assume there are 7 days in the week and the daily demand is split evenly throughout the week. For instance, if the weekly target is 700, then 100 per day. Make sure these calculation are rounded (hint: use the `ROUND` function). Finally, create another field called it **Meeting Daily Income**. This field takes two values: **Yes** if `Income >= Income Daily Target`; otherwise **No**.
- (e) (6 points) Add a **Aggregate** step to the resulting table form part (d). Aggregate the data based on **ProductName**, **Week**, and **Meeting Daily Income**. The aggregated table should look as the following:

	A	B	C	D	E
1	Number of Rows (Aggregated)	ProductName	Week	Meeting Daily Income	
2	7	Liquid Soap	28	NO	
3	4	Diffuser	17	NO	
4	3	Diffuser	13	NO	
5	4	Hand Wash	24	NO	
6	4	Hand Wash	10	NO	
7	3	Hand Wash	25	YES	
8	3	Hand Wash	28	YES	
9	2	Liquid Soap	7	YES	
10	2	Hand Wash	18	NO	
11	7	Bar Soap	13	NO	
12	1	Candle	31	YES	
13	3	Hand Wash	27	YES	
14	3	Candle	23	YES	
15	3	Hand Wash	32	YES	
16	1	Bar Soap	39	YES	
17	1	Bar Soap	33	YES	
18	3	Hand Wash	14	NO	
19	5	Diffuser	5	NO	
20	4	Diffuser	4	YES	
21	6	Bar Soap	25	NO	
22	5	Diffuser	27	YES	
23	6	Diffuser	3	NO	
24	6	Hand Wash	15	NO	
25	3	Hand Wash	37	YES	
26	3	Candle	24	YES	
27	7	Liquid Soap	22	NO	

Export the aggregated table to a `.csv` file called `product_week.csv`.

- (f) (6 points) Load the `product_week.csv` datafile to table desktop and duplicate the chart below.

## Products by Week



1

---

<sup>1</sup>Make sure you submit \*.tfl and \*.twb files in Blackboard.