

Request: Determine if there is a correlation between the weather and sales for the Underwater Basket Weaving division

Objective 1: Identify the goals of this request.

1. What are the actual questions bob is asking?
 - a. How do we define the weather?
 - i. States? Sunny, Cloudy, Rainy, etc.
 - ii. Metrics: temperature, humidity, wind, etc.
 - b. How granular is this investigation?
 - i. By location?
 - ii. Overall?
 - iii. Aggregated by date?
 - c. What type of sales are we considering?
 - d. What is the time period of analysis? Does Bob want information for the past week of sales, the last fiscal year, or is he looking for trends to use in projections? (I.e., if the weather next week is mostly sunny, how can Bob better predict the sales of baskets?)
2. What improvements to the business can potentially be achieved through this analysis?
 - a. Improved marketing campaigns?
 - b. More efficient management of staff time? (Number of sales people, for ex.?)
3. Is there empirical evidence that has lead Bob to want to investigate this relationship? Why are we doing this?
4. Are we willing to pay for additional data? Do we have a budget for this analysis?

Objective 2: Formulate a Plan

1. Identify potential sources for supplementary weather data
 - a. How do we rank different sources? What are our criteria for selection?
2. Document plan of analysis
3. Set schedule of deliverables, assign tasks and associated due dates to available staff to assist with analysis

Objective 3: Complete Analysis

1. Identify internal sales data
2. Join weather and transaction/sales data
 - a. Determine data pipeline/ETL structure
 - b. Where will this data be stored?
 - c. Use location of sale, date (and time, if available) of sale, and join that data with local weather data
3. Perform correlational analysis
 - a. Difference in mean sales by weather condition
 - i. One versus all comparison (if it rains, are sales significantly lower on average than under other weather conditions?)

- ii. Comparison of groups/Specific comparison (is there a significant difference in average sales across sunny and rainy days, for example?)
 - Adjusting for multiple hypothesis testing
 - iii. Depending on desired scope of investigation, consider comparison of mean sales by weather condition across specific geographies (such as county, city, state). Could also incorporate regional demographic data (census, etc.)
- b. Scatterplot/Correlation/Regression analysis of metrics such as temperature, wind speed, and precipitation versus total sales (per location and/or desired time period)
- 4. Data visualizations
 - a. Plot results to illustrate findings
 - b. What is the format of the charts?
 - c. Where will these be viewed?
- 5. Prepare brief presentation to share with Bob
 - a. Include visualizations as well as brief description of approach and overall findings from analysis
 - b. Include actionable recommendations based on patterns identified in the analysis; explain how these recommendations relate back to Bob's initial question

Objective 4:

1. Review presentation with project team
2. Address questions, revise accordingly

Objective 5:

1. Present findings to Bob
2. Discuss
3. Address questions and revise if necessary
4. Consider investigating any new questions that arise as a result of this discussion
5. Go over actionable recommendations, discuss implementation needs and plan with Bob and other relevant team members