

Take Home Questions

1. We often receive datasets from potential customers in various formats. This data is called Package Level Details (PLD) and helps us determine where the customer has the most demand.
 - a. Please load the provided dataset received directly from a potential faux customer ("202302 BIE Technical Interview PLD1.csv") and prepare it for analysis.

- OriginZip = Source of Packages
- DestinationZip = Zip of the Final Delivery Address of Packages

OrderCount	OriginZip	DestinationZip	Length	Width	Height	WeightOunces
1	12409	60197	12.0	9.0	1.0	6.0
2	49782	13144	9.0	8.0	6.0	24.0
1	42323	80920	12.0	9.0	1.0	6.0

- b. In order to make comparisons across customers and time ranges, we normalize the package demand; create a feature of the average number of packages per day from a given OriginZip to DestinationZip (assume this is data for the month of Sept 2022).
2. With the PLD, we often want to identify potential "markets" that would place our sortation capabilities nearest the demand to maximize our impact to the customer. For example, if we see tons of volume near Orlando, FL, we might recommend to the customer to send all that unsorted volume to a sortation facility in Orlando, and then sort to the final destination once closer to the destination. How would you identify an ideal area to set up a sortation facility near the destination zip codes - No code needed/Pseudocode Acceptable.
 3. Please load in the next dataset ("202302 BIE Technical Interview PLD2.csv"), from a different faux customer. Assume all packages come from a single fulfillment center at **33602**. How would you productionalize the above process to take in various different formats of customer PLDs and enable internal stakeholders, of varying levels of technical experience, to retrieve the results - No code needed/Pseudocode Acceptable.

shipmonth	shippingzipcode
3	07747
2	36471
2	49423