**Research Question:** Do people order different foods at different times of day? For example, do people order more junk food relative to other foods late at night?

**Technical Skills:**Explore the data using Pandas. Visualize the data to help address main research question. Tell a compelling story by presenting results in Jupyter notebook.

**Description:**Instacart is a grocery store delivery service. People place orders online, requested products are gathered at the grocery store, and delivered to the person’s location, sometimes within an hour. Instacart made data for several million of these orders publicly available on their website ([https://www.instacart.com/datasets/grocery-shopping-2017 (Links to an external site.)](https://www.instacart.com/datasets/grocery-shopping-2017)). I have summarized the data into an easy to read csv file (InstacartOrdersByDepartment.csv).

Your job is to explore the data with the goal of constructing a visualization that addresses the research question: Do people order different foods at different times of day? There are a few reasons we might expect products orders to be different for different times of day. First, some orders may be impulse purchases in which a person makes an order for things they want right away (e.g. frozen pizza late at night). Second, if a person is ordering groceries at a particular time of day some products may be come to mind more easily (e.g. thinking about coffee in the morning). A person may also be inspired to place an order because they ran out of a product they were just using at that time (e.g. milk in the morning). Instacart provided the full list of products per order. To make it easier to explore I have summarized taken a count of all products ordered per department in each hour of the day.

**Data Codebook:**Each row is a summary of orders associated with a particular hour in the day and a particular department.

* order\_hour\_of\_day: The hour of the day when the order was placed (0 = 12am, 1 = 1am etc.)
* department: The department the products came from (e.g. alcohol, produce).
* num\_orders\_hour: The number of products ordered from this particular department during this particular hour.
* Tot\_orders\_dept: The total number of products ordered from this particular department across all hours of the day.

**Deliverable:**

* Executable Jupyter notebook which presents your final calculations and visualizations. Notebook should include at least the following:
  + YOUR NAME
  + Citation for the source of data and explanation of the data
  + Final calculations and visualizations only (i.e. not ones you decided weren’t worth presenting).
  + Explanation of each step you took and why that would make sense to a layperson.
  + Explanation of the results and what they mean in terms of the research question.
  + Take-away message which succinctly makes it clear what was found and why it matters.
  + A note about the limitations of the data or analyses.
* The goal is to construct at least one visualization. This visualization should include at least the following:
  + Products bought by time of day.
  + Legend with intelligible labels.
  + Intelligible axis labels.

**Challenges:**

* Which is more informative raw counts of number of products ordered per department per hour or the proportion of products ordered per department per hour ?
* Does it make to include all departments? Or would it make sense to focus on a couple of departments?
* Does it make sense to visualize hours per day? Or should you group hours into categories (e.g. morning, evening)? If you wanted to use categories which would you choose and how would you make this decision?
* How could these results be actionable to a company?
* How could these results be actionable to a Instacart consumer?
* What other information might you want to know about the data set?
* Is this the right question to be asking? Is there another related question that would tell us something more interesting or relevant?