

# Assignment

Consider the bakery dataset:

*BreadBasket\_DMS.csv*

Using the enclosed script "transactions\_from\_bakery.py" we have added a few extra fields (such as period: night, morning, afternoon, evening). The expanded dataset is in file

*BreadBasket\_DMS\_output.csv*

You will write a Python script to read this data into Pandas and then write additional code to answer the following questions. Round your answers to 2 decimal digits

## Questions:

1. what is the busiest (in terms of number of transactions)?
  - (a) hour
  - (b) day of the week
  - (c) period
2. what is the most profitable time (in terms of revenue)?
  - (a) hour
  - (b) day of the week

(c) period

3. what is the most and least popular item?
4. assume one barrista can handle 50 transactions per day.  
How many barristas do you need for each day of the week?
5. divide all items in 3 groups (drinks, food, unknown). What  
is the average price of a drink and a food item?
6. does this coffee shop make more money from selling drinks  
or from selling food?
7. what are the top 5 most popular items for each day of the  
week? does this list stays the same from day to day?
8. what are the bottom 5 least popular items for each day of  
the week? does this list stays the same from day to day?
9. how many drinks are there per transaction?