**Data Annotation and Scripting:**

**PART I**

Attached is a file (included in the task) containing price data for hotels in New York. The prices for different hotels were obtained through multiple samples; each sample includes prices for different hotels on various dates, up to one month from the sampling date. Note that it is not guaranteed that all hotels will appear in every sample. The file contains the following fields:

a. Snapshot - the sample identifier for which we requested the prices (you can ignore this data).  
b. Snapshot Date - the sampling date.  
c. Checkin Date - the date of hotel check-in.  
d. Days - the duration of the stay in the hotel in days - note that you only received data for a 5-day stay.  
e. Original Price - room price without any special discounts - in dollars.  
f. Discount Price - room price after discount (the discount depends on the discount code type) - in dollars.  
g. Discount Code - discount code (a number between 1 and 4) - affects the discount amount.  
h. Available Rooms - the number of available rooms on this date (note that -1 indicates that information about the number of rooms was not available).  
i. Hotel Name - hotel name.  
j. Hotel Stars - the hotel's star rating.

Tasks to Perform:

Using Python code and Pandas, you need to add the following columns to the data file:  
a. DayDiff - the number of days between the Snapshot Date and the Checkin Date.  
b. WeekDay - the day of the week (Sun/Mon/Tue/Wed/Thu/Fri/Sat) of the Checkin Date.  
c. DiscountDiff - the discount amount in dollars (the difference between Original Price and Discount Price).  
d. DiscountPerc - the discount percentage.

The file after these changes should be named: Hotels\_data\_Changed.csv.  
  
**PART II**

For this task, feel free to choose an image by yourself.

Here are some links with images as examples:

* <http://images.cocodataset.org/val2017/000000296649.jpg>
* <http://images.cocodataset.org/val2017/000000184791.jpg>
* <http://images.cocodataset.org/val2017/000000502136.jpg>
* <http://images.cocodataset.org/val2017/000000386912.jpg>

1. **Image Annotation:**
   1. Manually create rectangular bounding boxes for two objects (as you wish) in the selected (single) image.
   2. Save the bounding box coordinates in JSON format. The JSON should look like this:

{

  "image\_name": "image1.jpg",

  "annotations": [

    {

      "label": "animal",

      "coordinates": {

        "x\_min": 50,

        "y\_min": 100,

        "x\_max": 200,

        "y\_max": 250

      }

    },

    {

      "label": "person",

      "coordinates": {

        "x\_min": 300,

        "y\_min": 350,

        "x\_max": 450,

        "y\_max": 500

      }

    }

  ]

}

1. **Scripting:**
   1. Write a Python script that takes the image and its JSON annotation file as input. The script should output a CSV file with normalized coordinates of the bounding boxes. The normalization should be done by dividing the coordinates by the image's dimensions (width and height).

The CSV should have the following format:

label, x\_min, y\_min, x\_max, y\_max

animal, 0.10, 0.20, 0.40, 0.50

person, 0.60, 0.70, 0.90, 1.00

* 1. Modify the Python script to also output another JSON file that converts the bounding box coordinates from **x\_min, y\_min, x\_max, y\_max** to x**\_min, y\_min, width, height**. The new JSON should look like this:

{

  "image\_name": "image1.jpg",

  "annotations": [

    {

      "label": "animal",

      "coordinates": {

        "x\_min": 50,

        "y\_min": 100,

        "width": 150,

        "height": 150

      }

    },

    {

      "label": "person",

      "coordinates": {

        "x\_min": 300,

        "y\_min": 350,

        "width": 150,

        "height": 150

      }

    }

  ]

}

* 1. To each one of the two steps above, create a visualization of the results.