Capital Bikeshare Project Data Wrangle Explanation

The Capital bikeshare datasets required little data wrangling other than renaming a few columns based on preference, formatting the date and time columns to match with the weather data. The 2017 hourly weather dataset was more challenging for various reasons outlined in question/answer format below:

1. **Renaming the columns as below:**

    'instant': 'rec_id',

    'dteday':'datetime',

    'holiday': 'is_holiday',

    'workingday': 'is_workingday',

    'weathersit': 'weather_condition',

    'hum': 'humidity',

    'mnth': 'month',

    'cnt': 'total_count',

    'hr': 'hour',

    'yr': 'year'

2. **There were not any missing values to drop or replace. Type casting the attributes as 'datetime' or 'category' shown below**

    stats['datetime'] = pd.to_datetime(stats.datetime)#dae time conversion

```
# categorical variables

stats['season'] = stats.season.astype('category')

stats['is_holiday'] = stats.is_holiday.astype('category')

stats['weekday'] = stats.weekday.astype('category')

stats['weather_condition'] = stats.weather_condition.astype('category')

stats['is_workingday'] = stats.is_workingday.astype('category')

stats['month'] = stats.month.astype('category')

stats['year'] = stats.year.astype('category')

stats['hour'] = stats.hour.astype('category')
```

3. Data Wrangling iPython Notebook link may be found here:

https://github.com/rashi-n/Machine-Learning-Projects/blob/master/Capstone%20Projects/CS%20I-%
20Data%20Wrangling.ipynb