

f) How does revenue vary over time? Calculate the average trip revenue per month - analysing it by hour of the day (day vs night) and the day of the week (weekday vs weekend).

```
from mrjob.job import MRJob
from datetime import datetime

class VendorDetails(MRJob):

    def get_dateformat(self, datetime_str):
        formats = ['%Y-%m-%d %H:%M', '%Y-%m-%d %H:%M:%S', '%d-%m-%Y %H:%M:%S', '%d-%m-%Y %H:%M']
        for fmt in formats:
            try:
                return datetime.strptime(datetime_str, fmt)
            except ValueError:
                pass
        raise ValueError('No valid date format available')

    def mapper(self, _, line):
        # Skip the header line as it contains header of csv
        if not line.startswith('VendorID'):
            fields = line.split(',')
            revenue = float(fields[10])
            pickup_datetime = self.get_dateformat(fields[1])
            month = pickup_datetime.month
            hour = pickup_datetime.hour
            # Consider time before 6 PM as day and after Night
            if hour > 18:
                flag = 'N'
            elif hour < 6:
                flag = 'N'
            else:
                flag = 'D'
            weekday = pickup_datetime.weekday()
            yield (month, flag, hour, weekday), revenue

    def reducer(self, key, values):
        total_revenue = 0
        num_trips = 0

        for revenue in values:
            total_revenue += revenue
            num_trips += 1

        average_revenue = total_revenue / num_trips

        yield key, average_revenue

if __name__ == '__main__':
    VendorDetails.run()
```