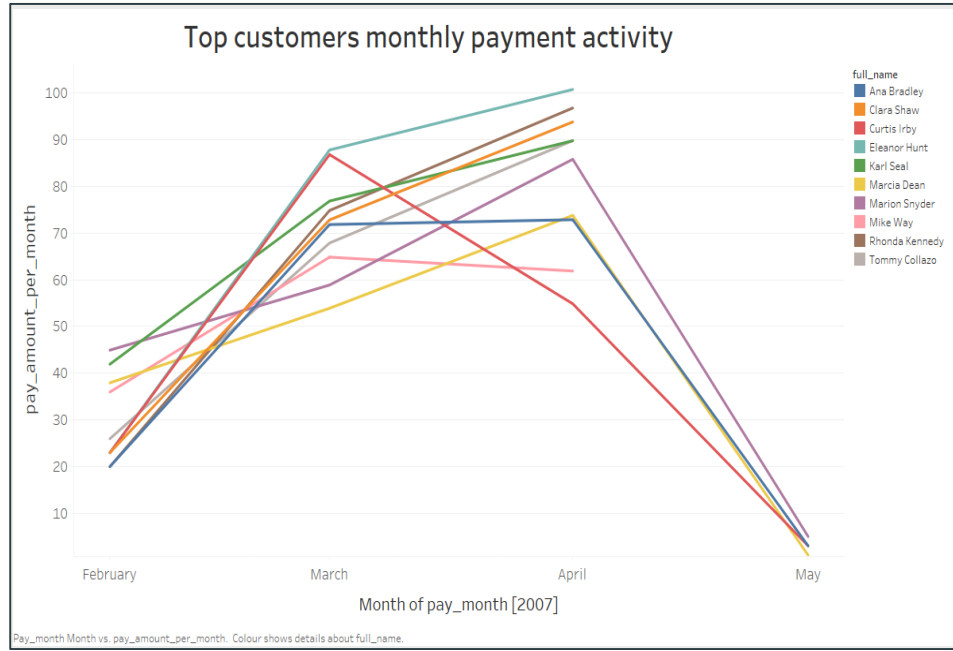# Programming for Data Science
with Python Nanodegree - Udacity
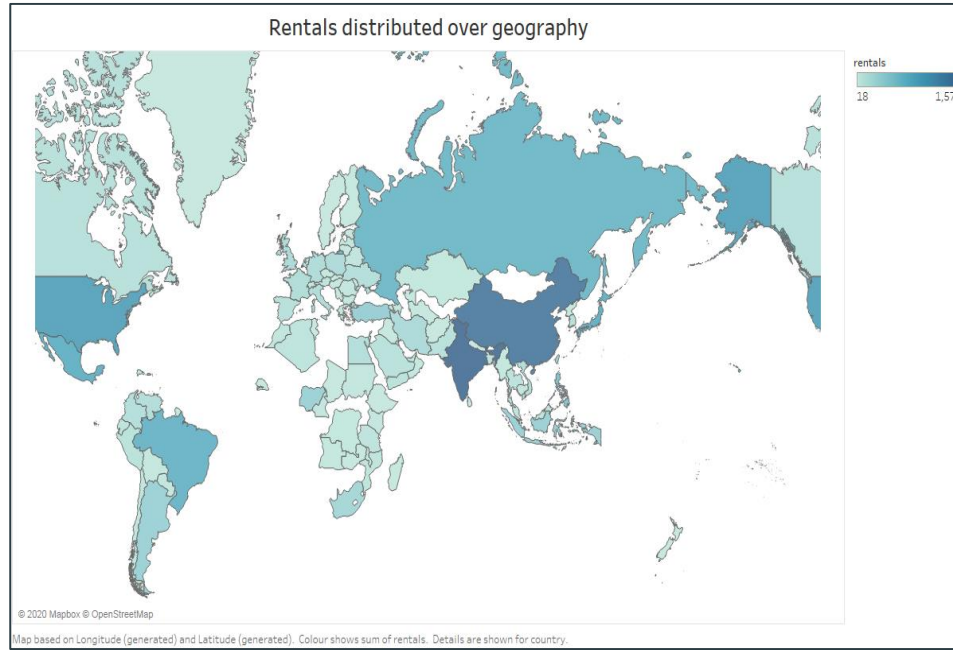
Postgre SQL Project:
Investigate a relational database (sakila movie database)

Q1: When does top customers paid most amount?


Top customers monthly payment activity

The highest payments from top customers are received in April, 2007 followed by March, 2007. However there is a significant drop in the payments after March, 2007. This might be the result of either excess payments from customers or lack of resources that meet the customers taste.

Rentals distributed over geography

rentals
18   1,572

© 2020 Mapbox © OpenStreetMap

Map based on Longitude (generated) and Latitude (generated). Colour shows sum of rentals. Details are shown for country.

The graph depicts the geographical spred out of the rentals over various countries. It can be observed that both India and China are the leading counties in terms of number of rentals, followed by Russia, United States and Brazil. Hence the customer accquisition costs must be cut in there regions and should be directed to other regions with low rentals and high population like Africa Canada and Europe.

Q3: Which films generate most income based on ratings?

Income generated based on film rating
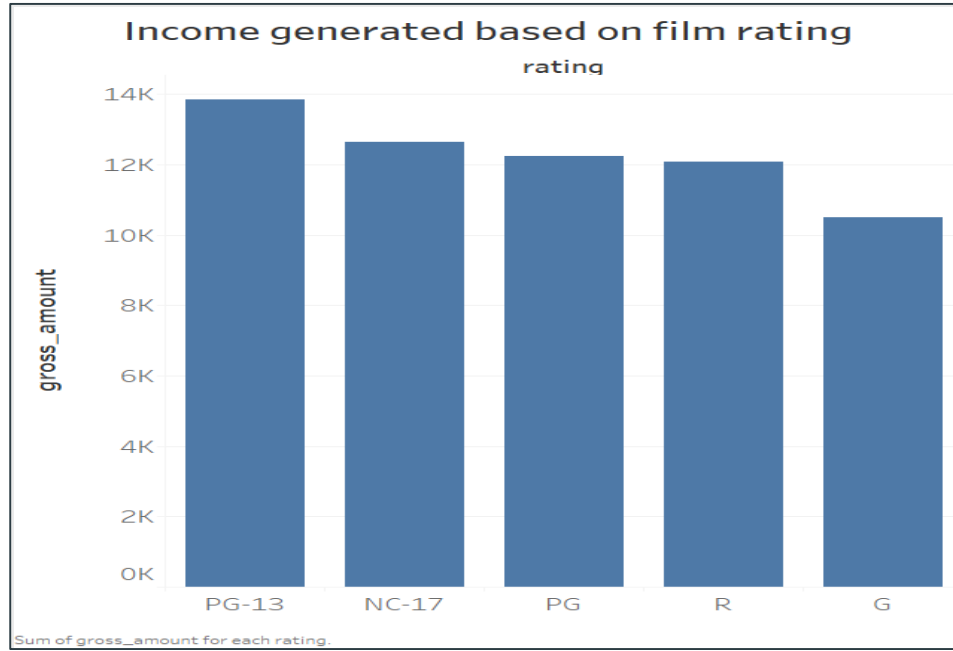
The films with PG-13 generated most income compared to films with different rating. However, it can be determined that the difference in income generated by films, based on the ratings are slim and can be overlooked. Since the ratings are not the most influencing factors in the income generation, other factors must have been playing a prominent role.

Q4: What are the most influencing factors in income generation?

Income based on film length & category
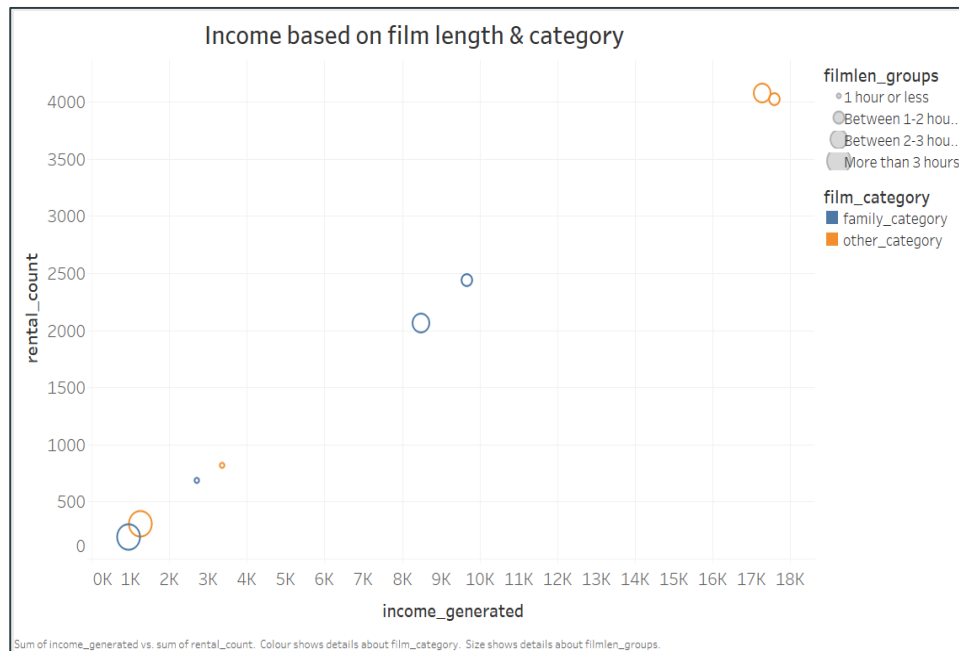
The graph depicts that the films with duration 'Between 1-2 hours' and 'Between 2-3 hours' are more likely to get rented by the customers. In turn, categories apart from 'family category', generate more income compared to 'family categories'.

This concludes that the film category and its duration plays a prominent role in income generation.

Q5: What influences the income generated by any actor?

Income generated based on Actors
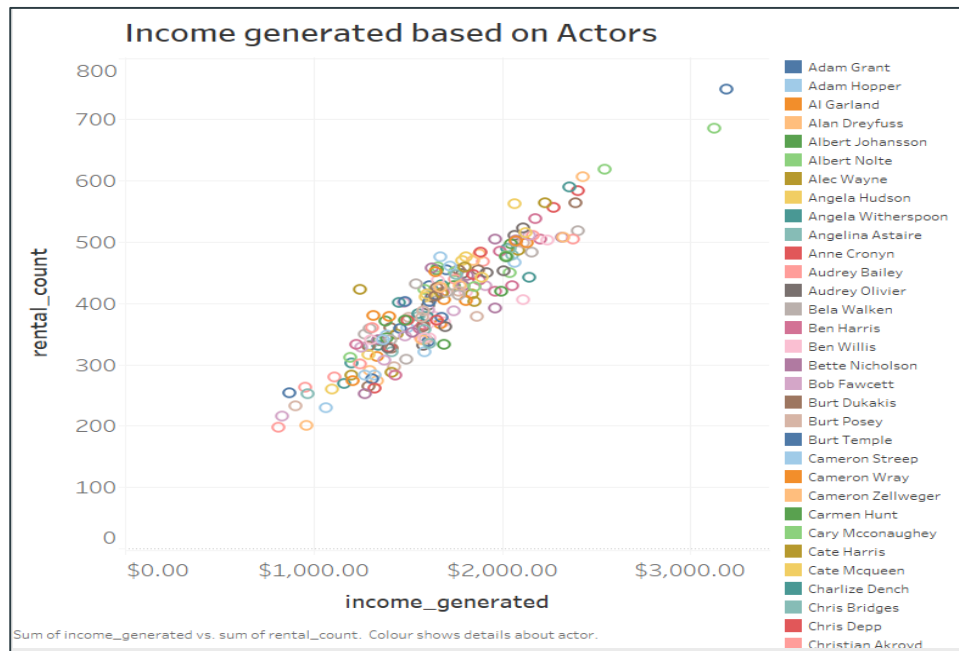
The graph reflects that the relation between, the number of times a dvd has been rented and income generated from it is a clear linear equation with a positive slope. This evaluates to the conclusion that the income generated by each actor is directly proportional to his popularity in having his films rented more frequent. This can be observed in the graph that some actors generate more income compared to less popular actors.

Q6: Who are the actors that generate most income ?

	actor text	rental_count bigint	income_generated numeric
1	Susan Davis	750	3193.49
2	Gina Degeneres	685	3129.17
3	Matthew Carrey	619	2543.78
4	Mary Keitel	607	2426.92
5	Scarlett Damon	519	2403.81
6	Walter Torn	584	2403.18
7	Henry Berry	564	2392.36
8	Christian Akroyd	504	2378.97
9	Angela Witherspoon	589	2357.11
10	Cameron Zellweger	507	2322.94

Data Output   Explain   Messages   Notifications

The resulting table is the list of top 10 actors that generate most income compared to other actors. Among them Susan Degeneres is the most popular actor generating a highest income of 3193.49 dollars. It can also be observed that some actors generate more income even when their films are rented comparatively less frequent. Scarlett Damon and Christian Akroyd are a good example for this occurrence.

Q7: Are there any overdues from the existing customers?

Customer data with overdues

customer_name	email	phone	movie	Year of rent..	return_date	overdue	
Adrian Clary	adrian.clary@sakilacusto..	182059202712	Shock Cabin	2006	Null	5075 days 04:24:29.2688..	Abc
Albert Crouse	albert.crouse@sakilacust..	256546485220	Varsity Trip	2006	Null	5075 days 04:24:29.2688..	Abc
Alberto Henning	alberto.henning@sakilac..	618156722572	Vanishing Rocky	2006	Null	5075 days 04:24:29.2688..	Abc
Alicia Mills	alicia.mills@sakilacustom..	761379480249	Sleepless Monsoon	2006	Null	5075 days 04:24:29.2688..	Abc
Allan Cornish	allan.cornish@sakilacust..	50898428626	Sundance Invasion	2006	Null	5075 days 04:24:29.2688..	Abc
Allen Butterfield	allen.butterfield@sakilac..	129953030512	Mulan Moon	2006	Null	5075 days 04:24:29.2688..	Abc
Allison Stanley	allison.stanley@sakilacus..	594319417514	Cyclone Family	2006	Null	5075 days 04:24:29.2688..	Abc
			Grapes Fury	2006	Null	5075 days 04:24:29.2688..	Abc
Ana Bradley	ana.bradley@sakilacusto..	72136330362	Happiness United	2006	Null	5075 days 04:24:29.2688..	Abc
Andy Vanhorn	andy.vanhorn@sakilacust..	995527378381	Tuxedo Mile	2006	Null	5075 days 04:24:29.2688..	Abc
Angela Hernandez	angela.hernandez@sakila..	18461860151	Canyon Stock	2006	Null	5075 days 04:24:29.2688..	Abc
Anna Hill	anna.hill@sakilacustomer..	911872220378	Boulevard Mob	2006	Null	5075 days 04:24:29.2688..	Abc
Annette Olson	annette.olson@sakilacus..	230250973122	Deer Virginian	2006	Null	5075 days 04:24:29.2688..	Abc
			Pirates Roxanne	2006	Null	5075 days 04:24:29.2688..	Abc
April Burns	april.burns@sakilacusto..	581174211853	Outfield Massacre	2006	Null	5075 days 04:24:29.2688..	Abc
Barry Lovelace	barry.lovelace@sakilacus..	689681677428	Creatures Shakespeare	2006	Null	5075 days 04:24:29.2688..	Abc
Becky Miles	becky.miles@sakilacusto..	648482415405	Virginian Pluto	2006	Null	5075 days 04:24:29.2688..	Abc
Bernard Colby	bernard.colby@sakilacus..	82088937724	Queen Luke	2006	Null	5075 days 04:24:29.2688..	Abc
Beth Franklin	beth.franklin@sakilacust..	470884141195	Baked Cleopatra	2006	Null	5075 days 04:24:29.2688..	Abc
Betty White	betty.white@sakilacusto..	517338314235	Groundhog Uncut	2006	Null	5075 days 04:24:29.2688..	Abc
Beverly Brooks	beverly.brooks@sakilacu..	427454485876	Shrek License	2006	Null	5075 days 04:24:29.2688..	Abc
Bill Gavin	bill.gavin@sakilacustome..	924663855568	Cleopatra Devil	2006	Null	5075 days 04:24:29.2688..	Abc
			Gleaming Jawbreaker	2006	Null	5075 days 04:24:29.2688..	Abc
Brandon Huey	brandon.huey@sakilacust..	99883471275	Ace Goldfinger	2006	Null	5075 days 04:24:29.2688..	Abc
Brent Harkins	brent.harkins@sakilacust..	854259976812	Lord Arizona	2006	Null	5075 days 04:24:29.2688..	Abc
Carmen Owens	carmen.owens@sakilacus..	272234298332	Affair Prejudice	2006	Null	5075 days 04:24:29.2688..	Abc
Carolyn Perez	carolyn.perez@sakilacust..	471675840679	River Outlaw	2006	Null	5075 days 04:24:29.2688..	Abc
			Torque Bound	2006	Null	5075 days 04:24:29.2688..	Abc

A total of 180 overdues were detected from the existing customers. The graph diaplays the list of customers with overdues along with their contact information and data related to their rentals as in the film rented, rental date and overdue count in days. The contact information of the customers is highly useful in the retrival of the rented dvd's.

Q8: When does the family films rented most based on the day of the week?

Rentals of family films based on day of the week
dow

Sum of rental_count for each dow.
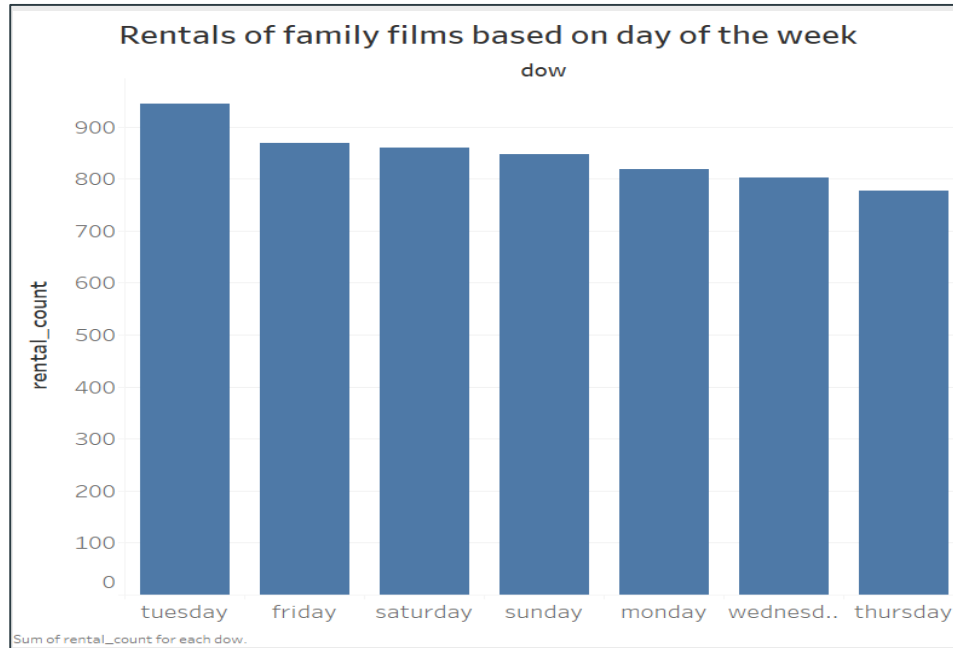
(Family films include below categories : Animation, Children, Classics, Comedy, Family and Music)

From the graph, it can be observed that the rentals of family films increase at the beginning of weekend and decline as the days go by. It can also be observed that the most rentals occur on the Tuesday despite not being the weekend.