**LOOQBOX**

**CHALLENGE SOLUTION**

Name: Matheus Nobre Gomes

E-mail: [matt-gomes@live.com](mailto:matt-gomes@live.com)

Phone: (11) 2471-7235

Cell phone: (11) 99566-0126 / (11) 96952-5592 (WhatsApp)

**SQL test – Solutions**

*Observations: You can see script.sql and other files that I used to more details*

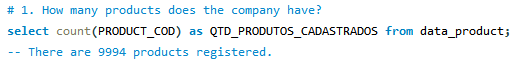
1. **How many products does the company have?**

The company has 9994 products registered.

Result Grid:



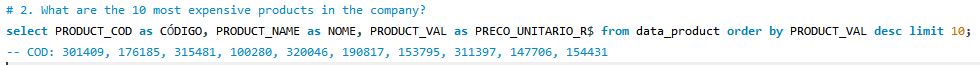
Command:



1. What are the 10 most expensive products in the company?

**Result Grid:**   

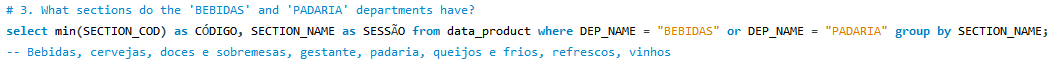

**Command:**



1. What sections do the 'BEBIDAS' and 'PADARIA' departments have?

**Result Grid:**  

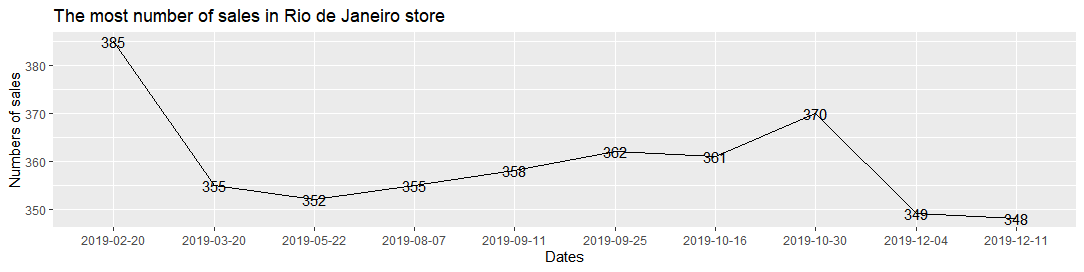

**Command:**



1. When were the most products sold? In which store?

|  |  |
| --- | --- |
| **Result Grid:** | **Command:** |

**Chart:**

****

1. What was the total sale of the products of each business area in the first quarter of 2019?

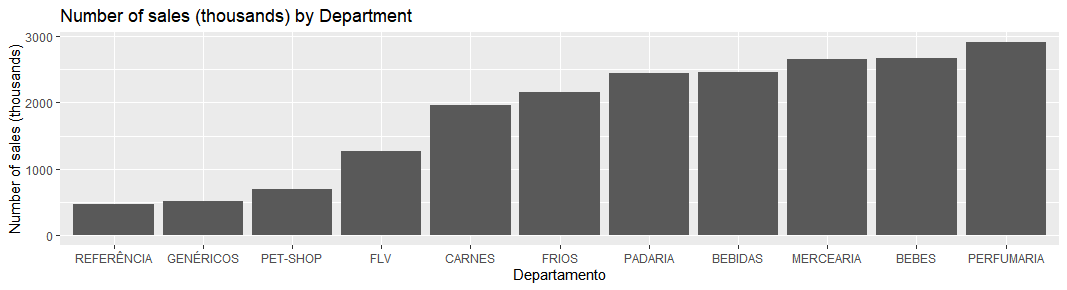
**Total:** 20.196.315 sales considering only sales that have a department registered.

|  |  |
| --- | --- |
| **Result Grid (1):** | **Command:** |

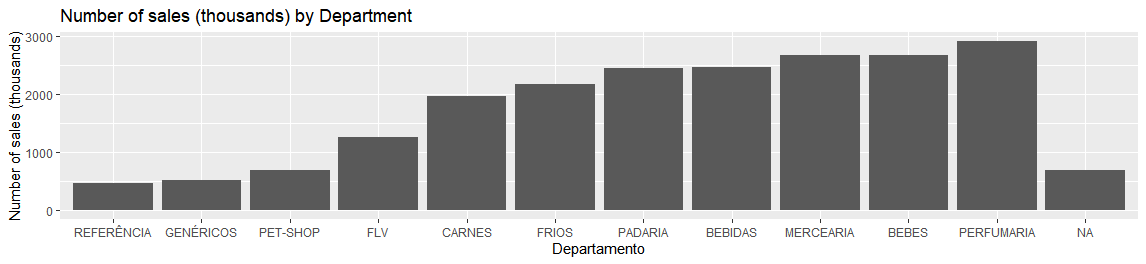
**Total:** 20.884.965 sales if considering sales that haven’t a department registered yet.

|  |  |
| --- | --- |
| **Result Grid (2):** | **Command:** |

**Chart (1):**



**Chart (2):**

****

**Building your own visualization**

1. Create at least one chart using the table IMDB\_movies. The code must be in R or Python, and you are free to use any libraries, data in the table and graphic format. Explain why you chose the visualization you are submitting

**Introduction**

I love watch movies in my leisure and my favorite genre is horror but the most of my friends don’t like it. Always when we’re looking for a horror movie, usually I already watched them and when I read this question and saw all the information you passed I thought:

* Now is the time to figure out why they don’t produce as many horror movies as the other genres.

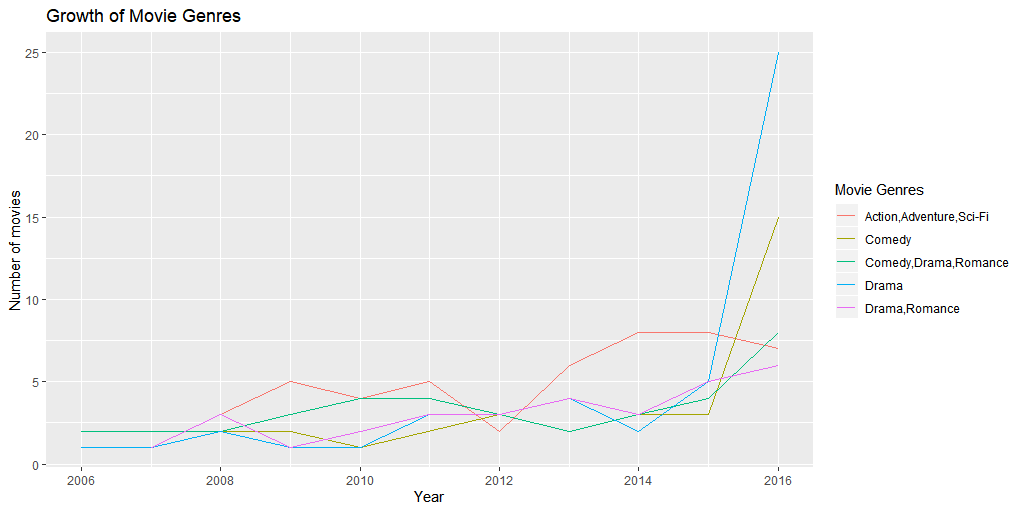
So, the objective of my visualization is understanding what the directors and the companies see when they are choosing a movie to produce.

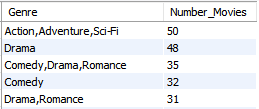
***Notes:***

* ***The distribution of movies per year is not proportional, so the analysis will not be so assertive.***
* ***There are 1000 samples among 2006 and 2016 and it’s a good quantity but if we add more samples will be better to my analysis.***

**Development**

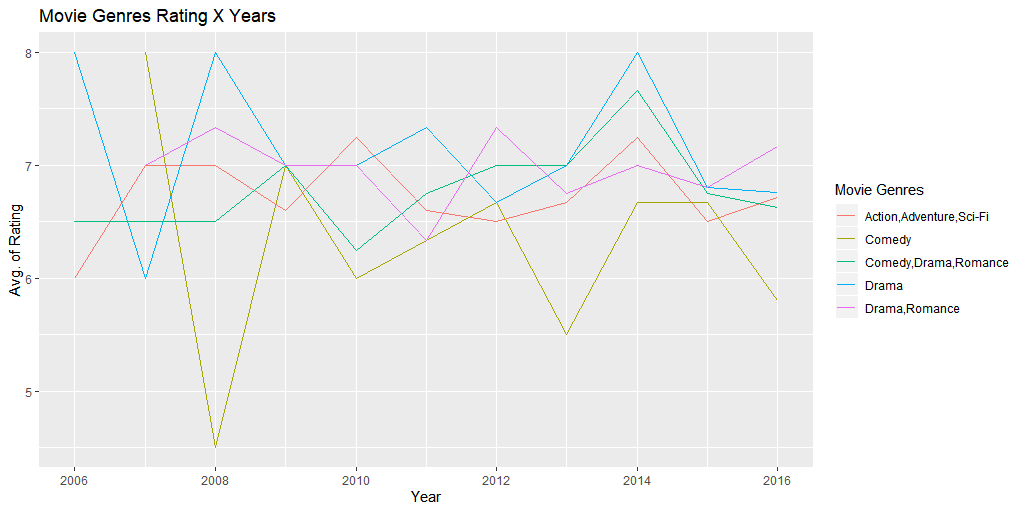
I figured out which genres have more movies for unpolluted chart and plotted one chart by year to see how these genres are behaved.



**Results:**

* The blue and Yellow has been growing exponentially since 2015 but in the past few years have maintained their average.
* The others genres have been keeping their average.

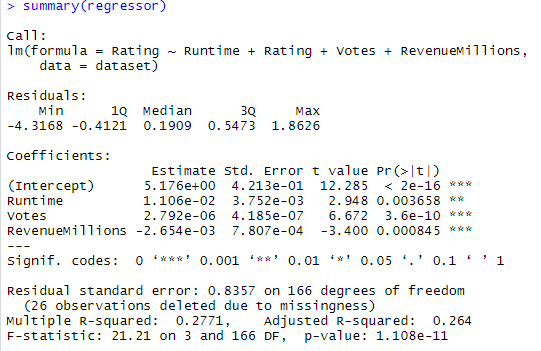
The next step was comparing their ratings and see if they have been behaving the same.



**Results:**

* The blue and Yellow didn't grow like their productions
* The others genres have been keeping their average.

**Conclusion:**

Analyzing this information, we can deduce the genre’s rating isn’t what decide if more movies will be produce because the rating doesn’t up or down like his productions. I need to get more information about these movies to compare, however I used multiple linear regression and backward elimination to see which the information is more significant and they are: runtime, votes and revenue.

I could analyze other information like which genres get more money or something like this but is more interesting for movie lovers like me understand this information as what is really important when you will produce a movie? What is the ideal runtime?

The first movie produced was at December 28th, 1895 and has since evolved. Today we have wonderful effects, 3d movies, 4d theaters and it makes me curious what theaters will look like in a few years.