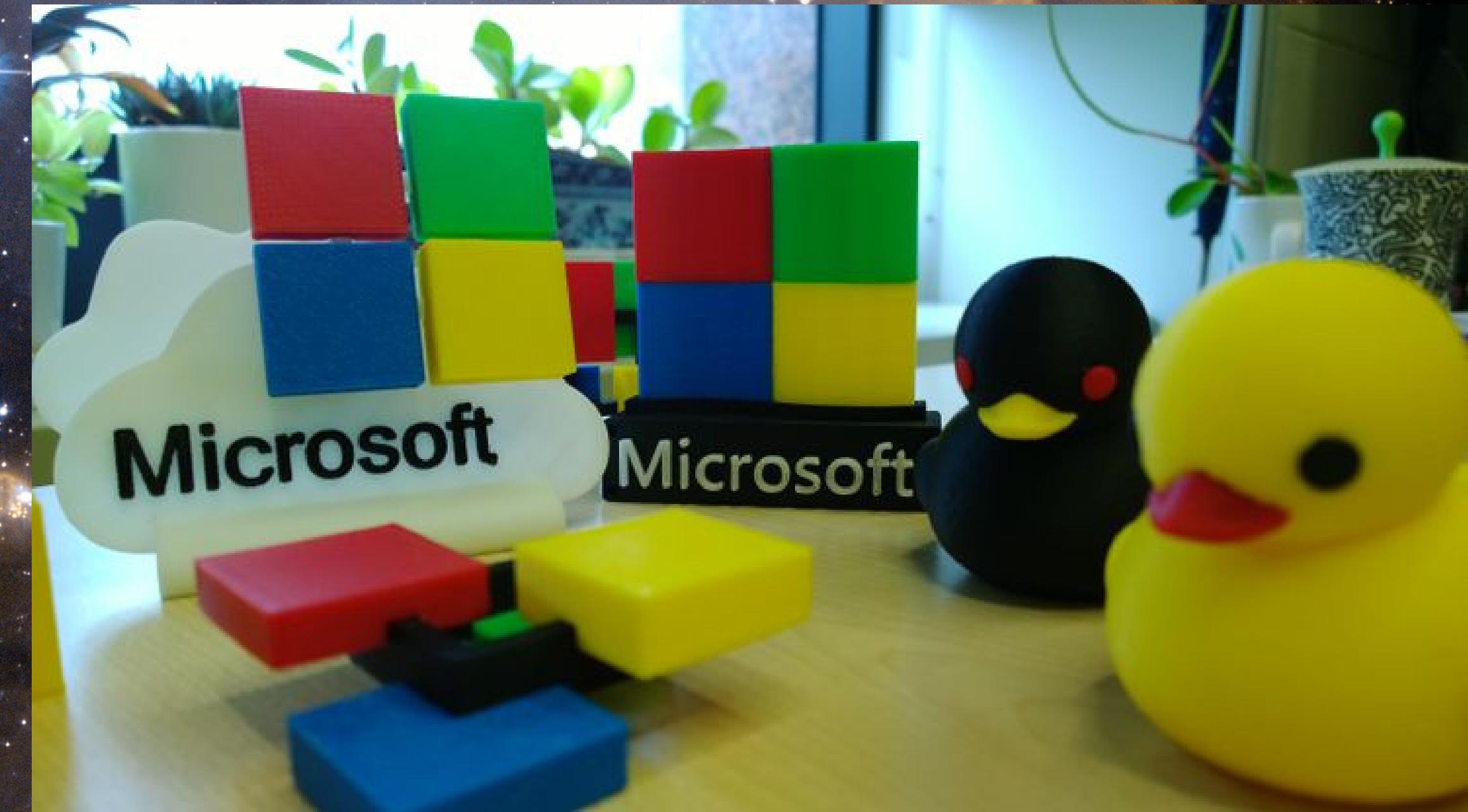


MICROSOFT MOVIE VENTURE



BUSINESS UNDERSTANDING

A good businessperson is one who is constantly looking for a market niche to fill or a growing company to invest in. Whatever the reason, investing in a new company involves risk and expense, so for an investor, the true question is always: Is the risk worth taking?

Microsoft wants to start a movie studio, but they know very little about the movie business. Therefore, this project does an exploratory data analysis on information from three datasets to assist Microsoft in learning more about the movie industry and determining whether the venture is worthwhile.

.

PROBLEM STATEMENT

The major issue is deciding whether or not to invest in the new business idea, as well as offering guidance on the aspects to consider before making the investment. What are the variables that are causing other film-related businesses to succeed so admirably?

PROPOSED SOLUTION

To present an analysis that will help microsoft decide whether or not the movie studio is a profitable investment.

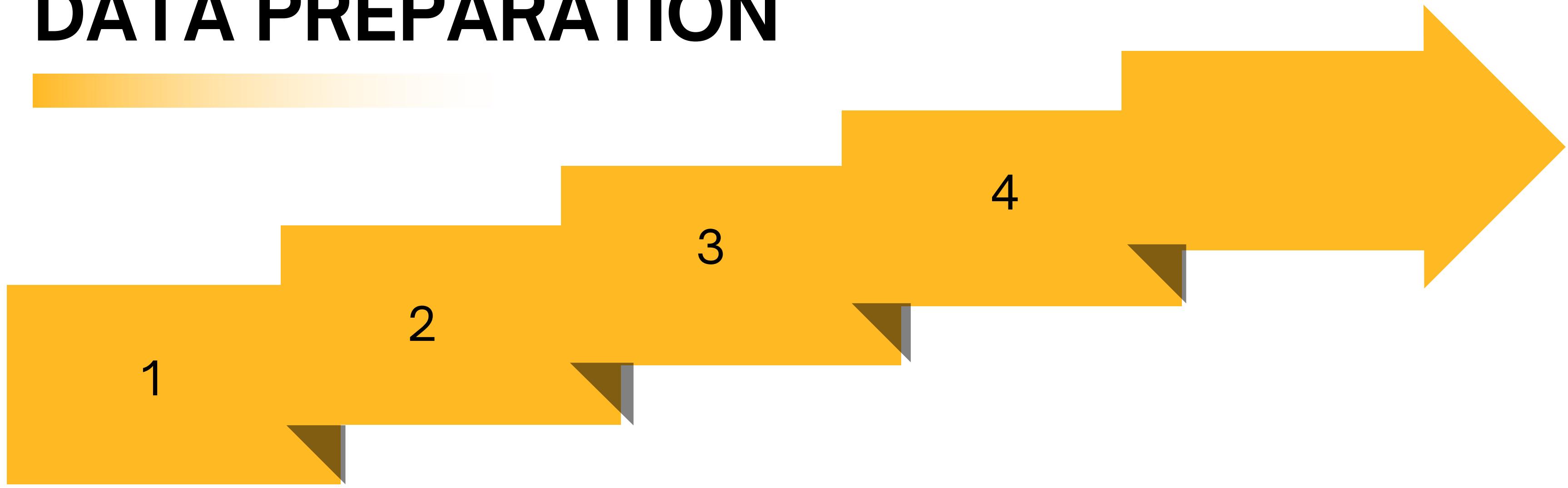
OBJECTIVES

To evaluate the expected production budget

To identify which Genres to invest in

To approximate the runtime for the movie in minutes

DATA PREPARATION



Loading

Data

Importing
libraries,readin
g and loading
of dataset

Data

Cleaning

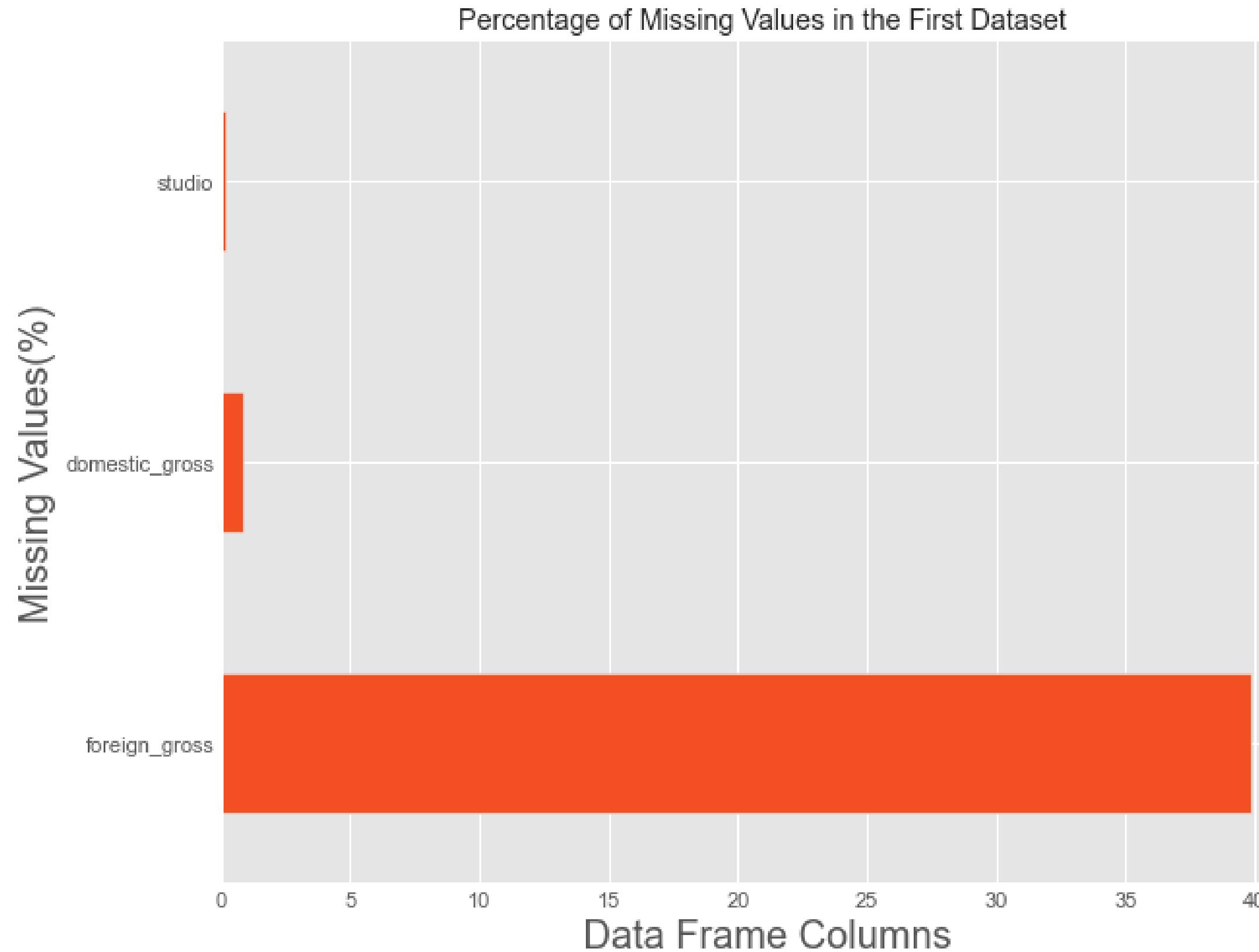
Removing null
and duplicate
values and
checking for
outliers,anoma
lies

Data

Analysis

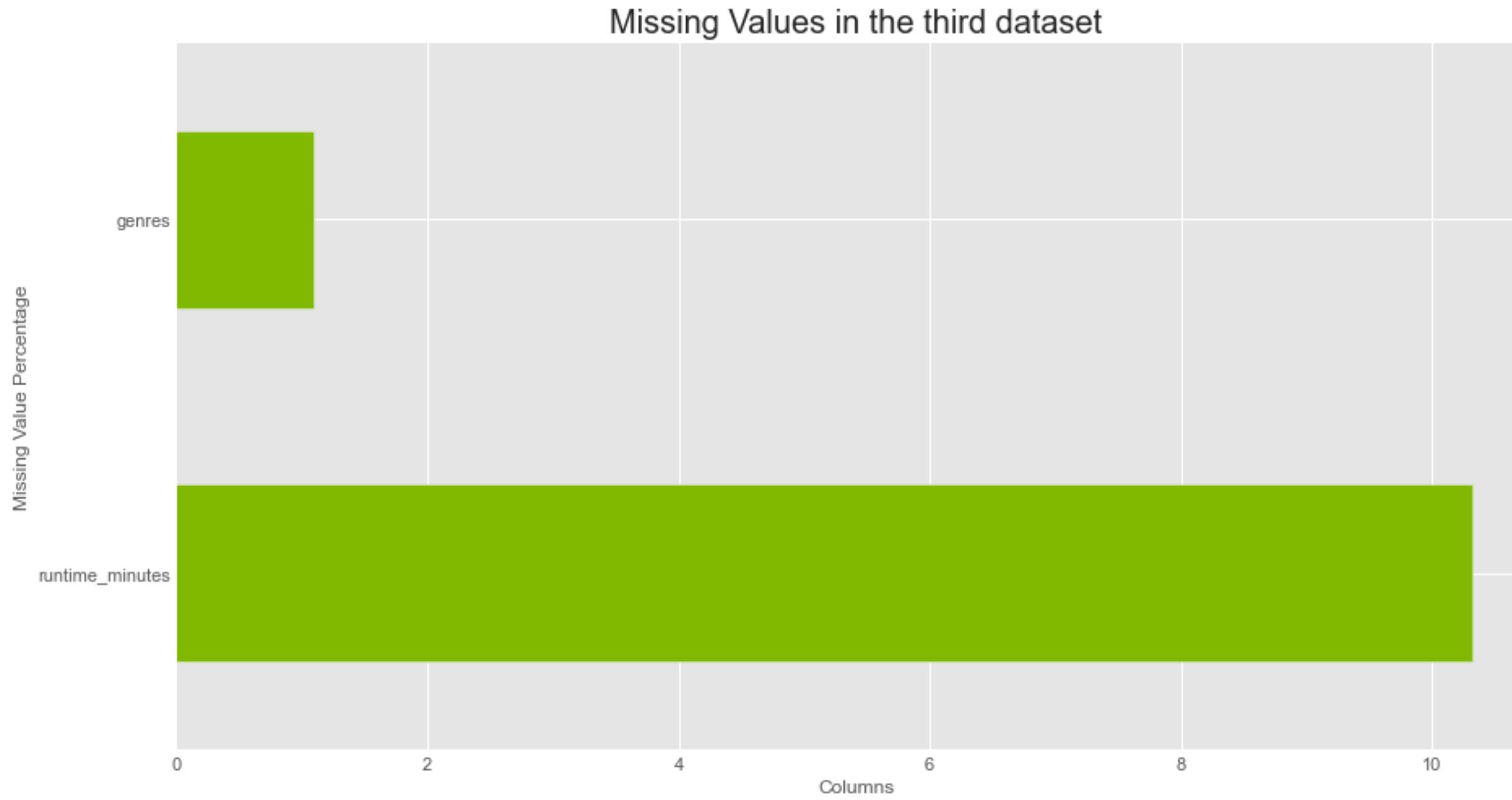
Visualizing how
different
variables
relate to each
other and to
the target
variable

MISSING VALUES IN THE FIRST DATASET



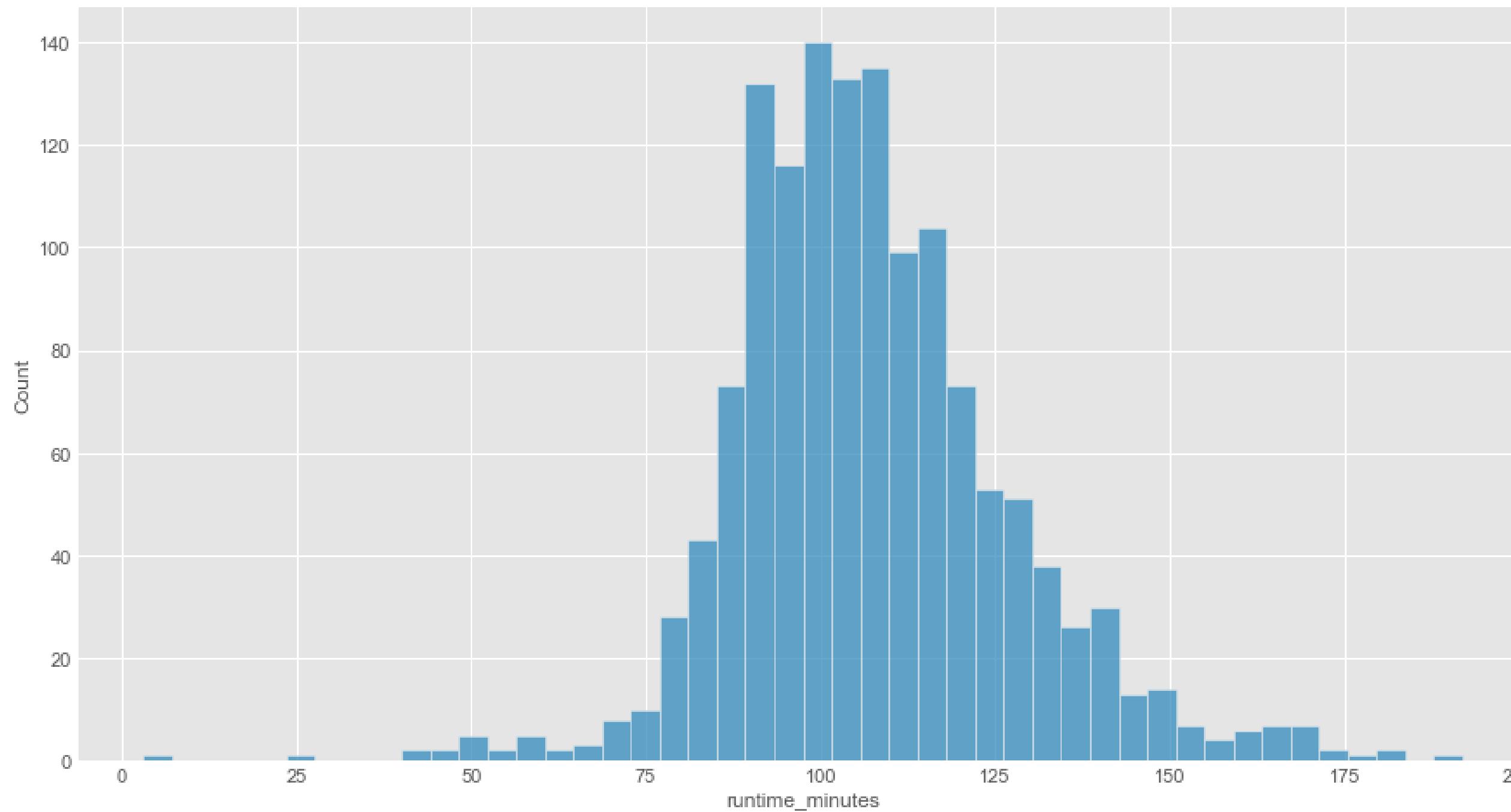
In the first dataset
the foreign_gross
had the highest
number of missing
values-1350

MISSING VALUES IN THE THIRD DATASET



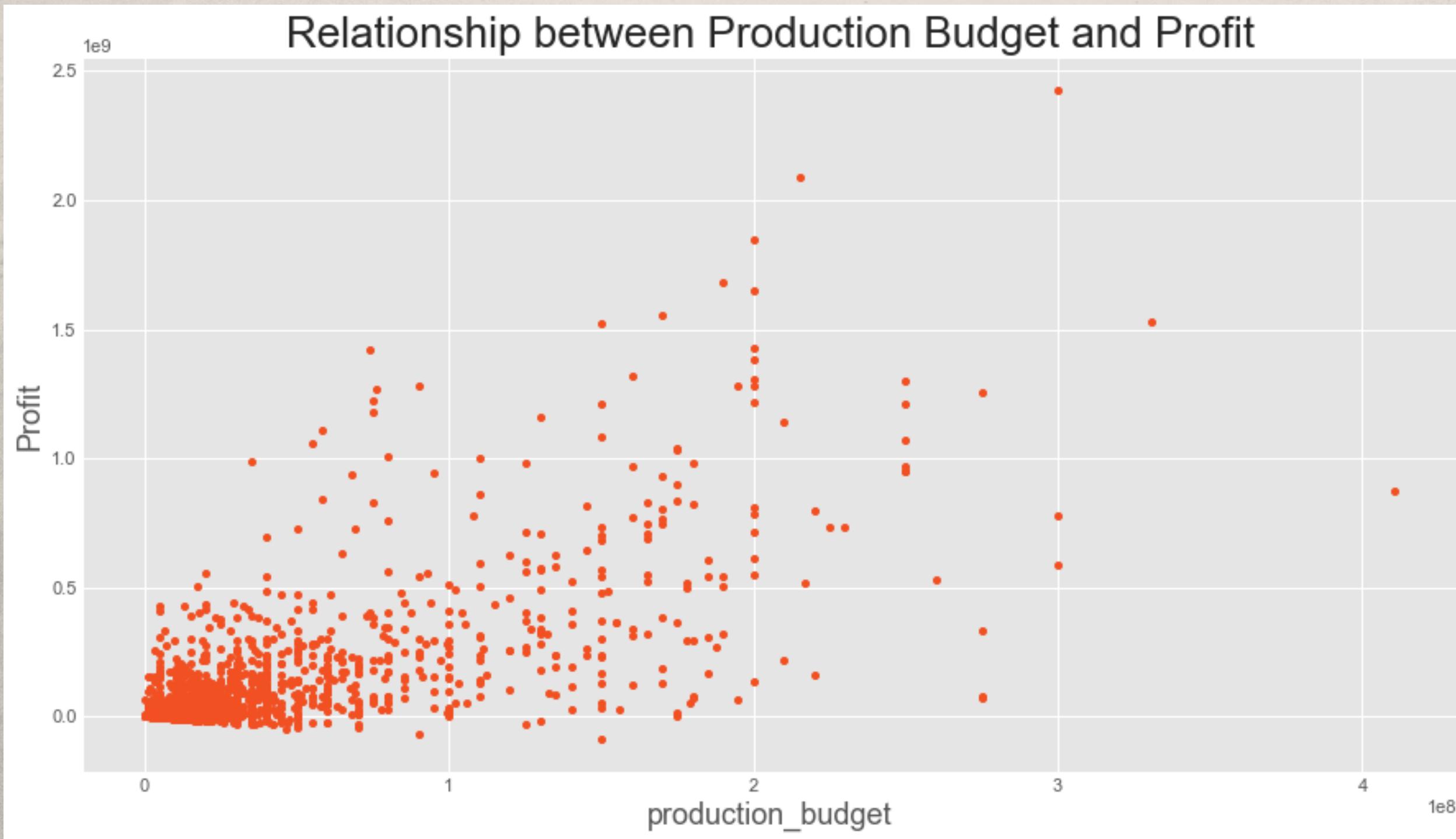
The runtime_minutes had the highest number of missing values (7620) which accounted for 10% of the total dataset.

FREQUENCY DISTRIBUTION IN THE RUNTIME IN MINUTES



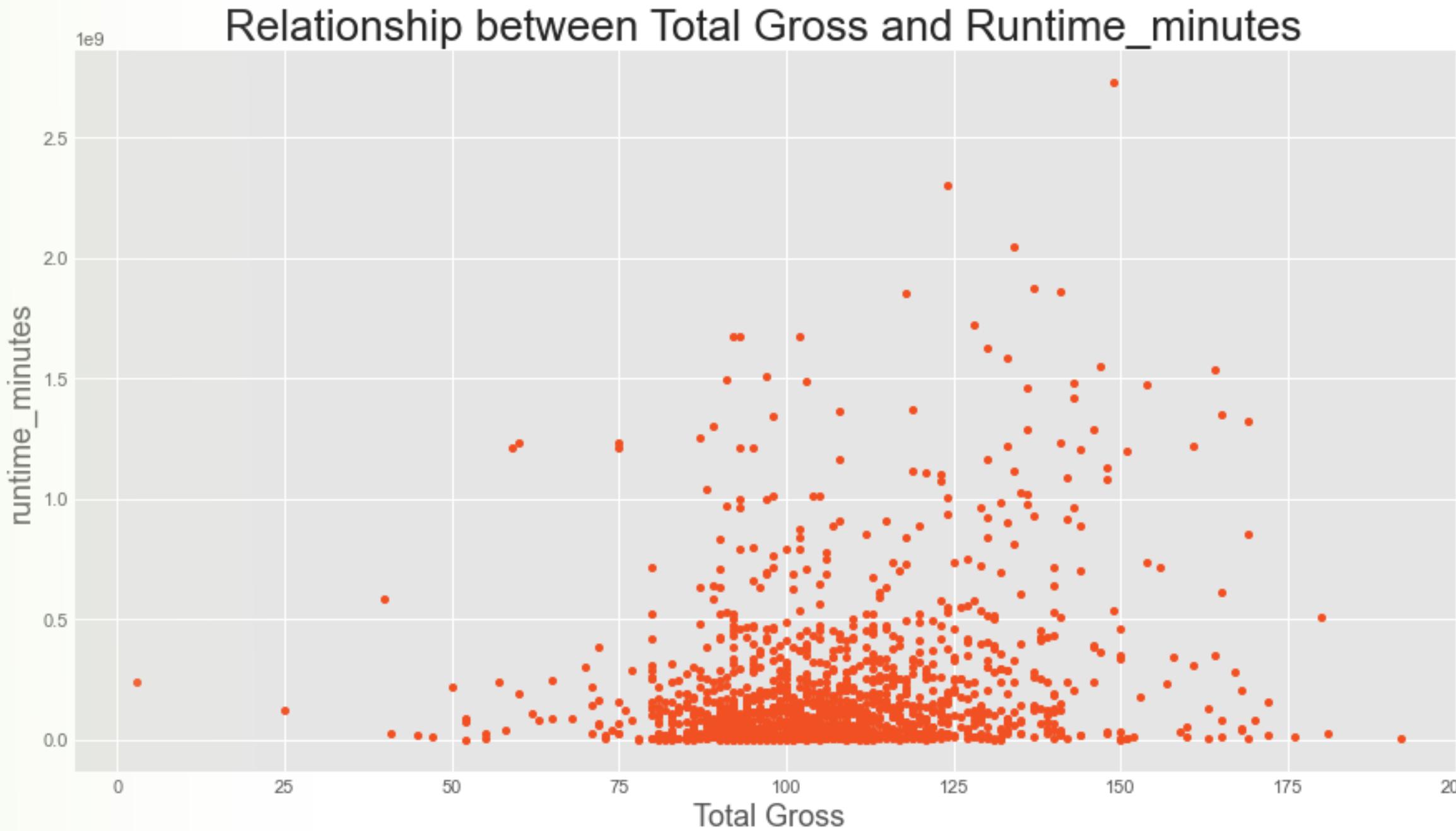
Most of the values are populated towards the centre meaning the runtime data had a normal distribution.

CORRELATION BETWEEN VARIABLES



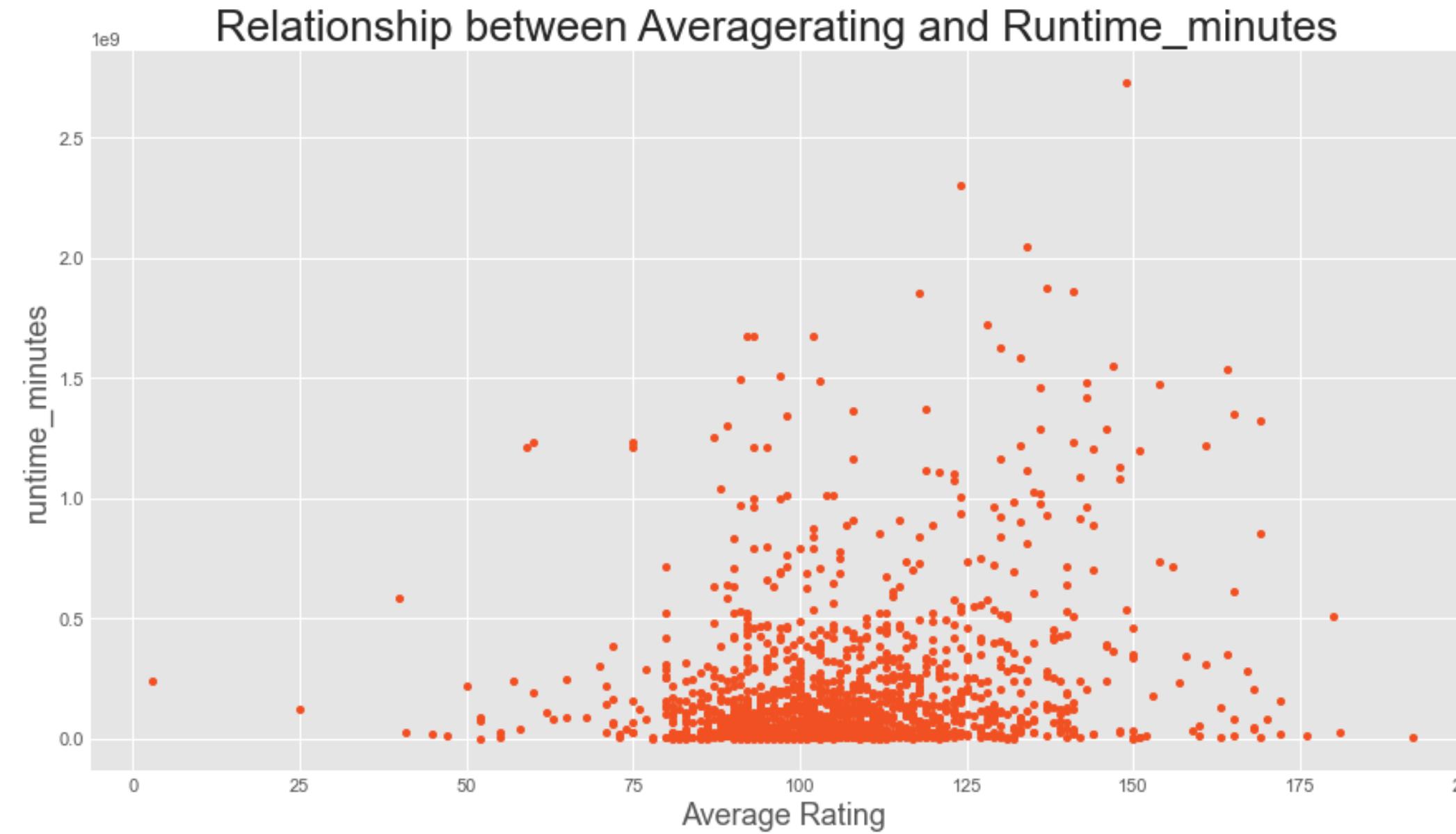
There is a strong positive correlation of 0.6771 between production budget and profit.

CORRELATION BETWEEN VARIABLES



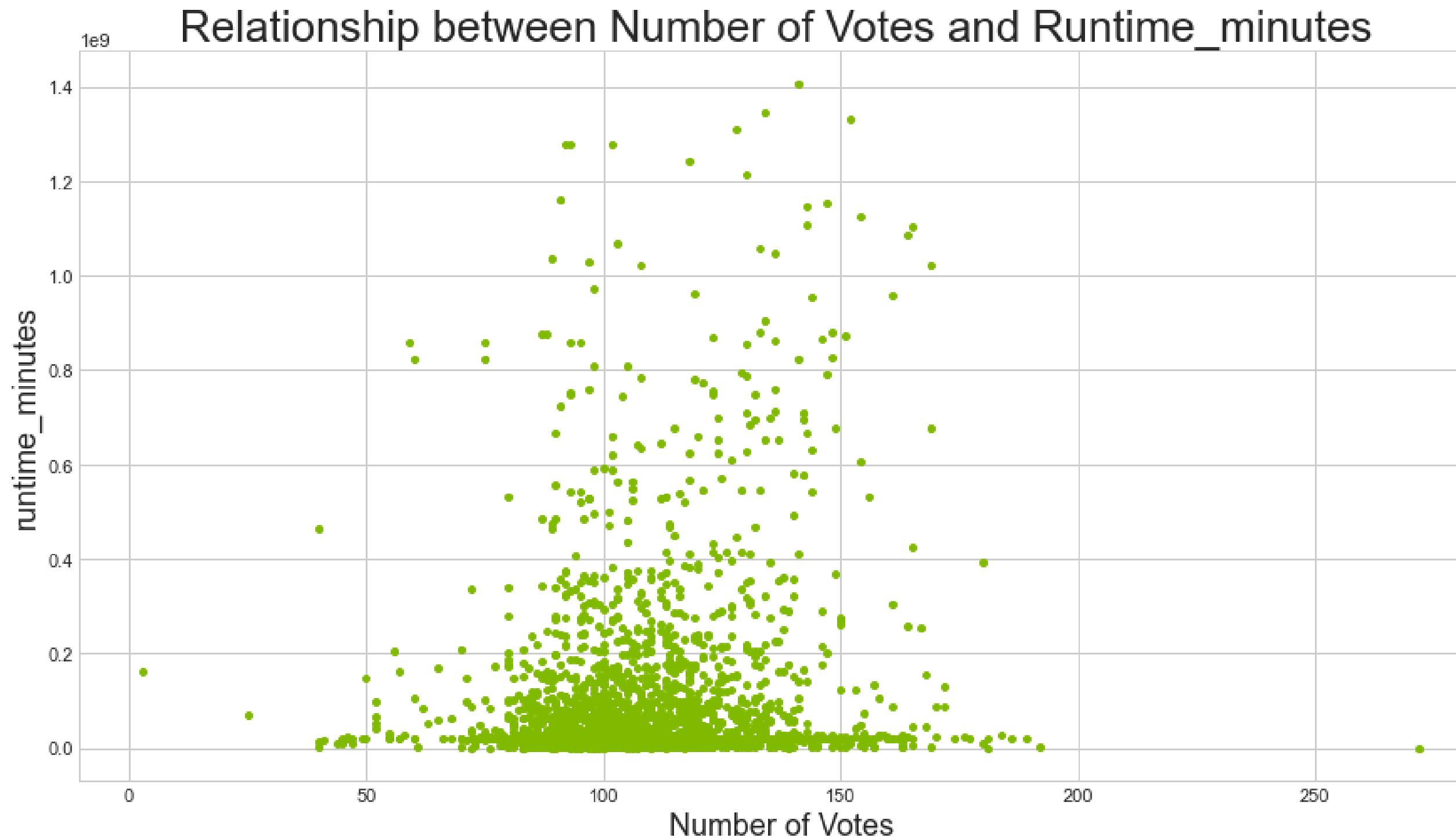
There exists a weak positive correlation of 0.1278
between total gross and runtime_minutes

CORRELATION BETWEEN VARIABLES



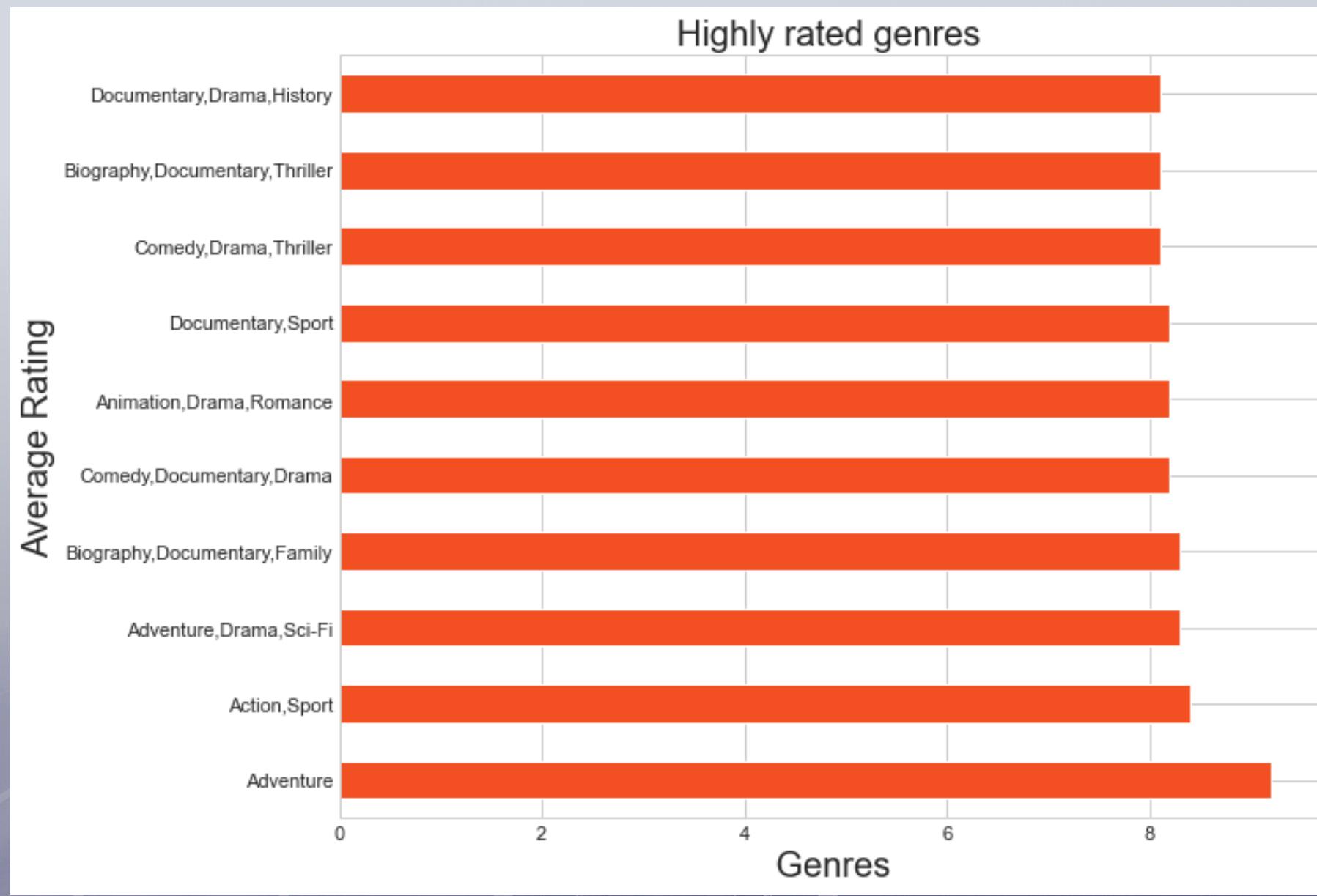
There is a weak positive correlation of 0.1222 between runtime in minutes and average rating.

CORRELATION BETWEEN VARIABLES



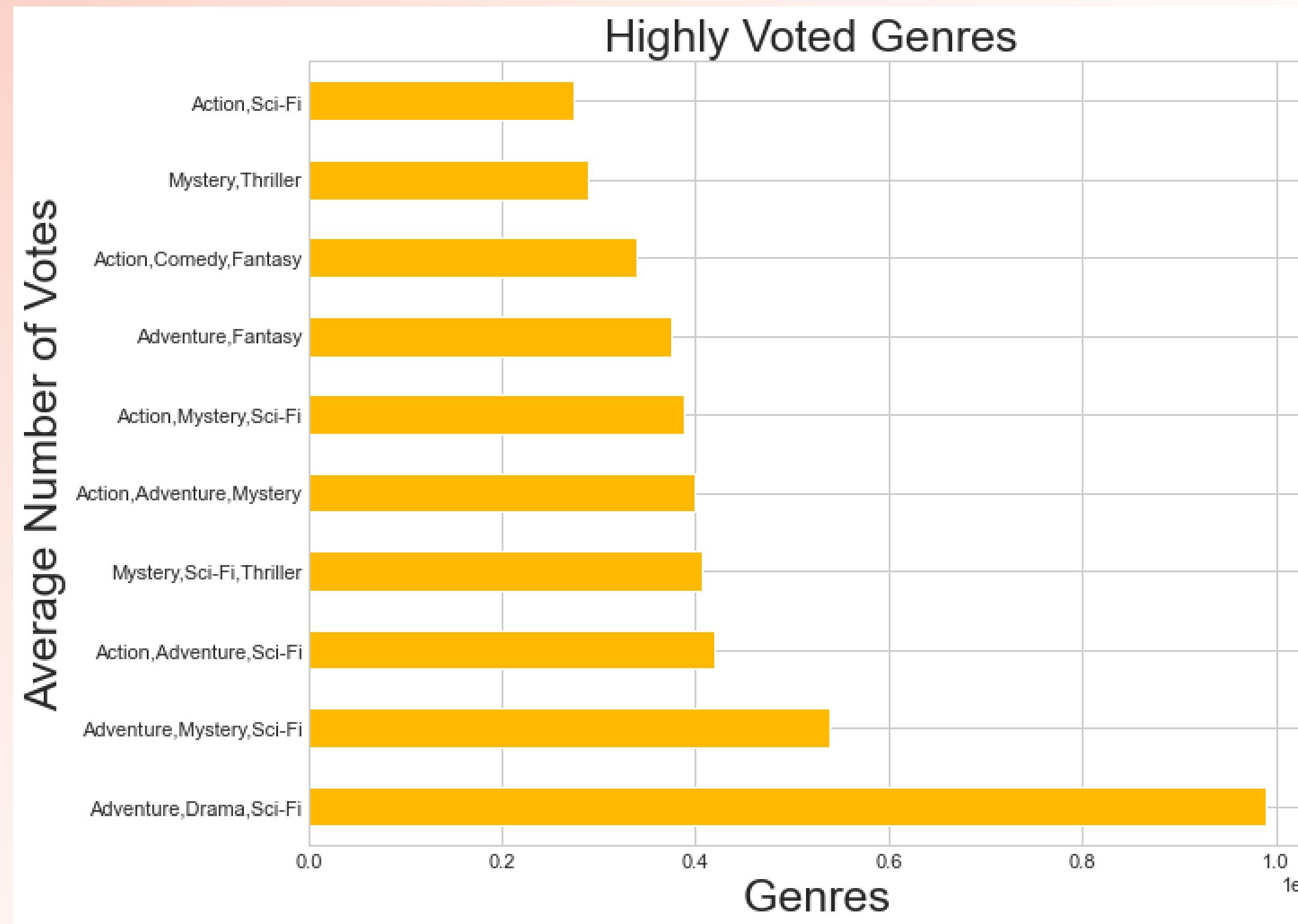
There is a very weak positive correlation of 0.2598 between runtime in minutes and the number of votes

HIGHLY RATED GENRES



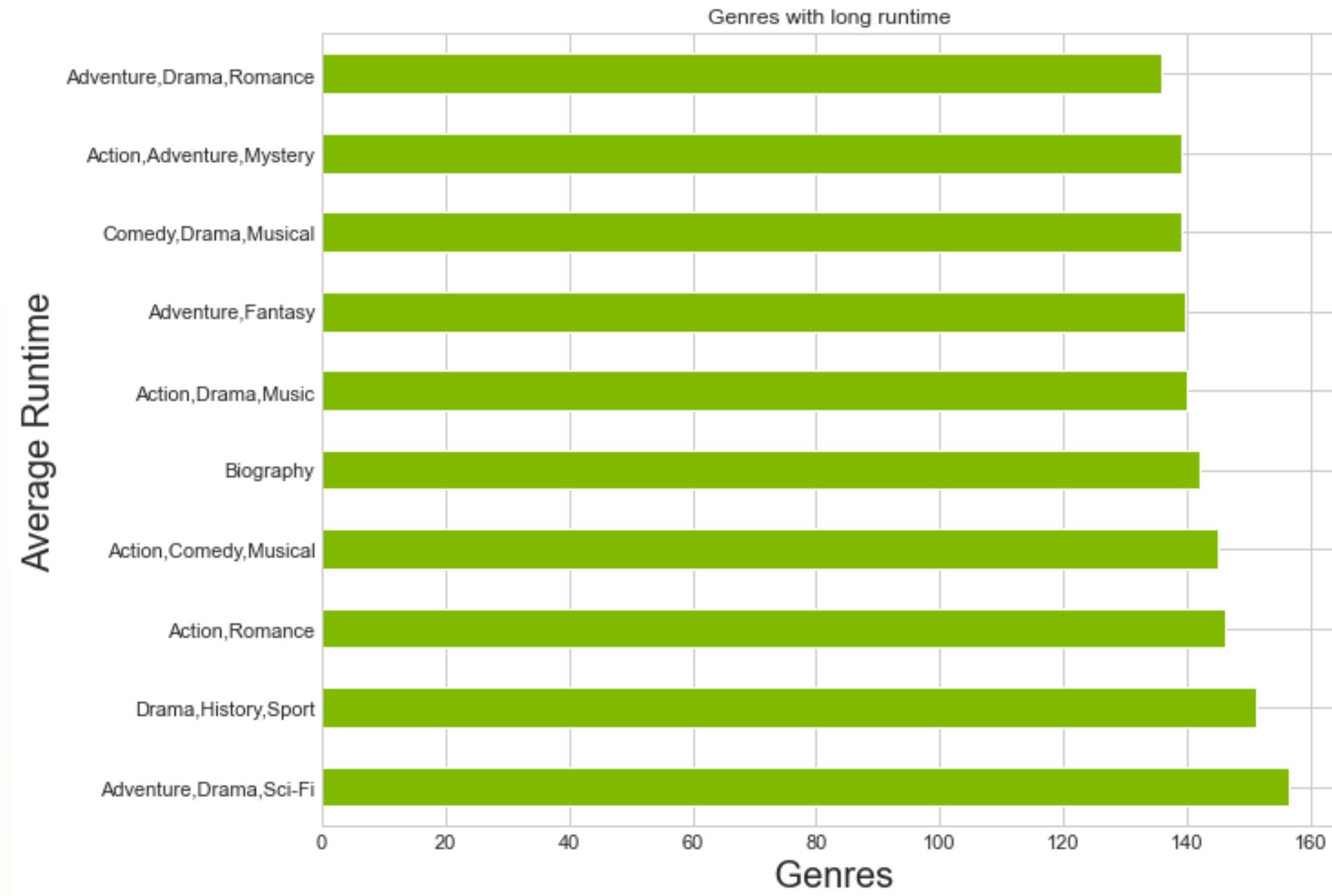
The adventure genre had the highest average rating (9.2)

MOST VOTED GENRES



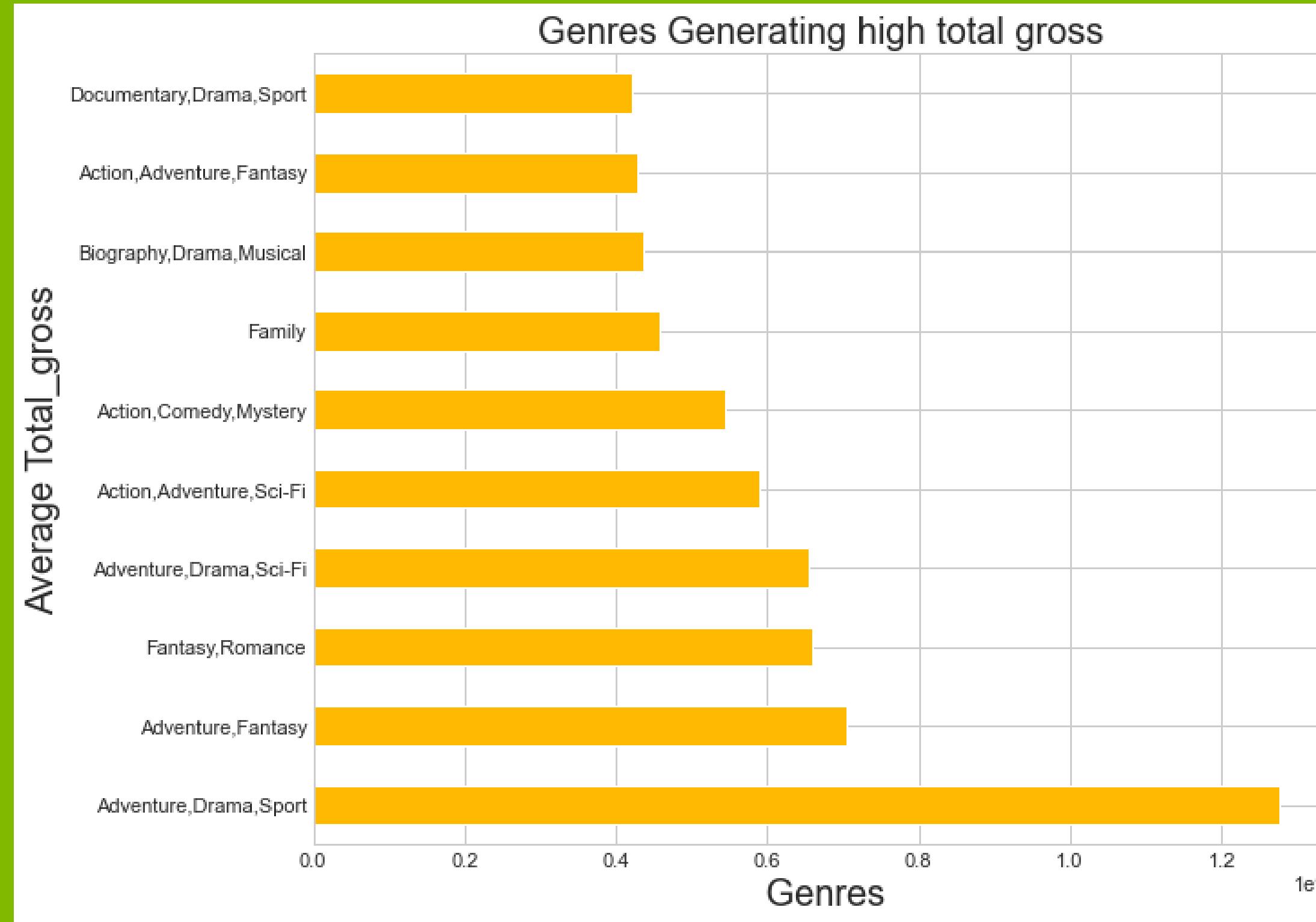
The adventure drama sci-Fi received the highest votes (989725)

GENRES WITH LONG RUNTIMES



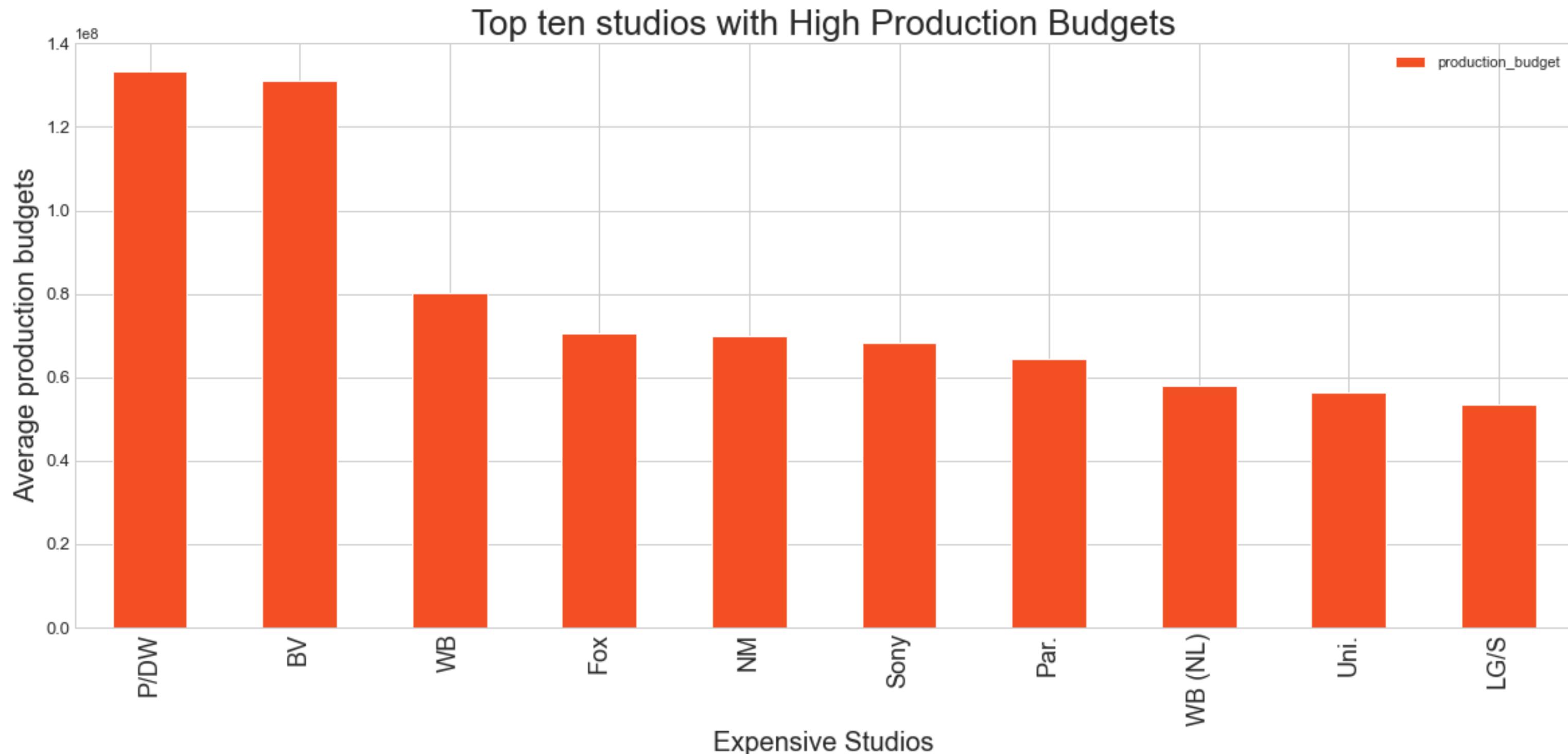
Adventure drama
Sci-Fi had the
longest average
runtime of 156.5
minutes.

GENRES GENERATING HIGH GROSS



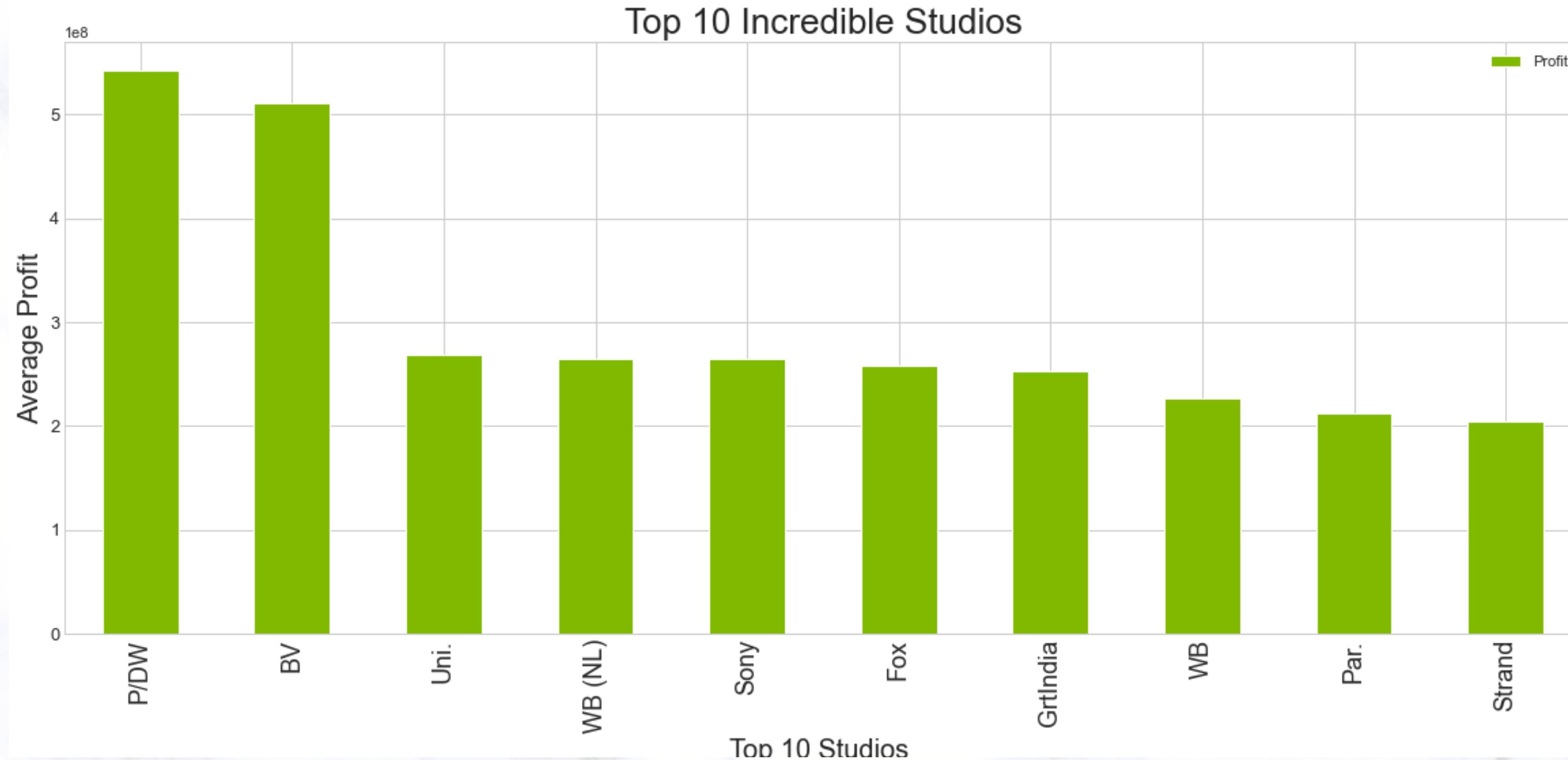
The adventure drama sport genre had the highest total gross of 1,276,400,000 dollars.

STUDIOS WITH HIGH PRODUCTION BUDGETS



P/DW studio has the highest production budgets of 133,400,000.

STUDIOS MAKING HIGHEST PROFITS



P/DW is the highest earning studio with a profit of 542,694,200 dollars.

CONCLUSION

Major factors to consider when choosing movie genre: Number of votes received, the average genre rating and most importantly total gross.

Criteria for measuring runtime in minutes should be the mean as the data is symmetrically distributed. Though weakly correlated with number of votes ,rating and total gross, this positive relations imply the longer the movie the better.

The P/DW studio had the highest profit and production budget hence the higher the production budge the higher the returns expected. Since it is skewed positively the median will be appropriate measure for approximation.

RECOMMENDATION



Genre

To invest in Adventure, Drama and Sci-Fi.

Production Budget

To set aside 25,000,000 dollars as the minimum for the production budget.

Movie Runtime

To have 107 minutes as the least runtime for the movie.

We welcome your



Name: Obwoge Getrude

LinkedIn: <https://www.linkedin.com/in/getrude-obwoge-59a1b21bb>