Precalc AI Lab

Modeling yearly movies by genre using exponential functions

Image generated with AI (Nano banana).

In this Lab you will write Python code and run it. You can write the code yourself or get an LLM to do it. Run the code in Google Colab. If it does not run correctly, rewrite until it works correctly.

Purpose of Lab

- Fit an exponential model of the form ab^t to data that exhibits approximate exponential growth.
- Given an exponential model, identify the values of a and b and relate them to the growth rate, growth factor, and starting value in context.
- Understand how a and b affect the graph.
- Compare parameters a and b in different models and interpret in context.

Context for Lab

Beginning in the 1980s, worldwide movie production increased approximately exponentially, likely driven by population growth and the home-video revolution (VHS, DVDs, rental stores). The trend continued until the early 2010s, after which the growth slowed, probably because of streaming platforms and the increased popularity of series. Different genres displayed different exponential growth rates.

To do this lab you need access to the data of the number of movies released by genre between 1980 and 2014. The file `movies_by_genre_1980_2014.csv` contains this data for several genres[1].

Part 1: Creating a Graph and Finding Equations

Let's consider the yearly releases for movies whose genre is "Romance". Start by loading the data into Colab.

Use an LLM

Generate Python code to graph the number of romance movies released each year from 1980 to 2014 and find an exponential function $f(t) = ab^t$ to fit the data. We ask the LLM to:

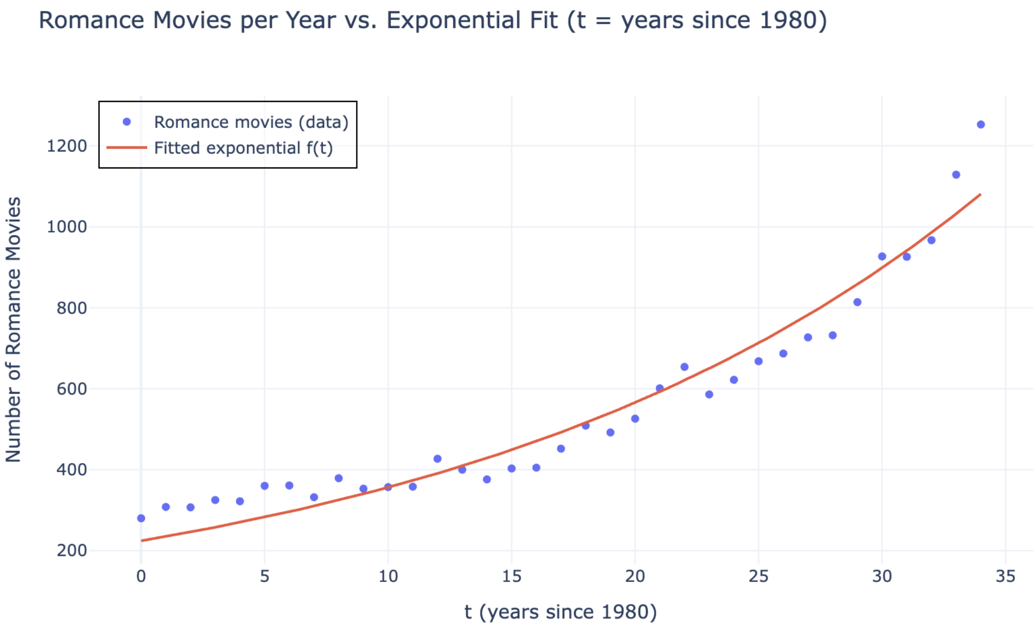
- Write simple, Colab-compatible code that is easy for a Precalculus student to follow.
- Use t in years as the independent variable, with $t = 0$ corresponding to 1980.
- Fit an exponential function to the data (directly, without using logs)..

[1] Data compiled from IMDB, `https://www.imdb.com/search/title`, accessed October 1, 2025.

- Plot the data together with the exponential model. Use the Plotly library for a more interactive plot.
- Use t on the horizontal axis (not the year).
- Print the equation $f(t) = ab^t$ for the exponential model.

Note: Let the LLM know the structure of the file `movies_by_genre_1980_2014.csv` (for example, by pasting in the contents or uploading the file) before requesting the code to reduce the chance that the code you get produces errors. In particular, be careful with capitalization in your code: the file `movies_by_genre_1980_2014.csv` has genre columns headings that are capitalized, like "Thriller" or "Sci-Fi".

We get the following graph (code at the end of the lab).



For the number of romance movies released each year since 1980, we get the exponential model

$$f(t) = 224.2222(1.047358)^t$$

Answer the following questions:

1. Evaluate $f(10)$. What year does that correspond to? What does the model predict for the number of Romance movies in that year? Is it close to the real value?
2. Evaluate $f(11)$. According to the model, what was the percentage increase in romance movies between 1990 and 1991?
3. Can you see a relationship between the your answer to the previous question and the value 1.047358 in the formula for $f(t)$?
4. Without any computation, what do you think the model predicts for the percentage increase in the number of Romance movies from 1991 to 1992? Check your answer.
5. Pick any other two consecutive years. What does the model predict for the percentage growth?
6. The model predicts $f(0) = 224.2222$. What does this mean in terms of romance movies? Does that look same as the number of romance movies that were released in 1980?

Part 2: Using and Interpreting Exponential Models

For an exponential function $f(t) = ab^t$:

- a gives the initial value (at $t = 0$).
- b is the growth/decay factor: $b > 1$ gives exponential growth, $b < 1$ gives exponential decay.

- $b = 1 + r$ where r is the percent change per unit of the input in decimal form (so 51% means $r = 0.51$).

Examples (assuming t is in years):

- For $g(t) = 3(1.14)^t$, we start (at $t = 0$) with $g(0) = 3$, and we have $b = 1.14$, so $r = 0.14$ which means 14%. The function $g(t)$ is growing 14% per year.
- For $h(t) = 2.3(0.98)^t$, we start (at $t = 0$) with $h(0) = 2.3$, and we have $b = 0.98$, so $r = -0.02$ which means -2%. The function $h(t)$ is shrinking 2% per year.

Answer the following questions about your exponential model for the number of romance movies released each year since 1980.

1. What is the value of a? What does it mean in terms of romance movies?
2. What is the vertical intercept of the graph of $f(t)$? What does the vertical intercept mean in terms of romance movies? How does it relate to the parameters a and b?
3. What is the value b? If you know the number of romance movies in a year, how can you use the value of b to predict the number of romance movies the next year?
4. By what percentage were romance movie productions increasing each year?

Part 3: Build another exponential model

Select another genre in `movies_by_genre_1980_2014.csv`, find an exponential model to fit the data, and graph it together with the data, either by editing the code or by using an LLM to generate new code. The exponential model you get may not be as good as for romance but should show an approximate exponential trend.

Use your graph and exponential model to answer the following questions:

1. Compare the values of the parameter a for the two genres. What does it tell you about the differences between them?
2. Compare the values of the parameter b. What does it tell you about the differences between the genres?
3. By what percentage was the number of movies produced for your selected genre increasing each year?

How different are the graphs of the exponential models for your genre and for the romance genre? Plotting them on the same graph helps.

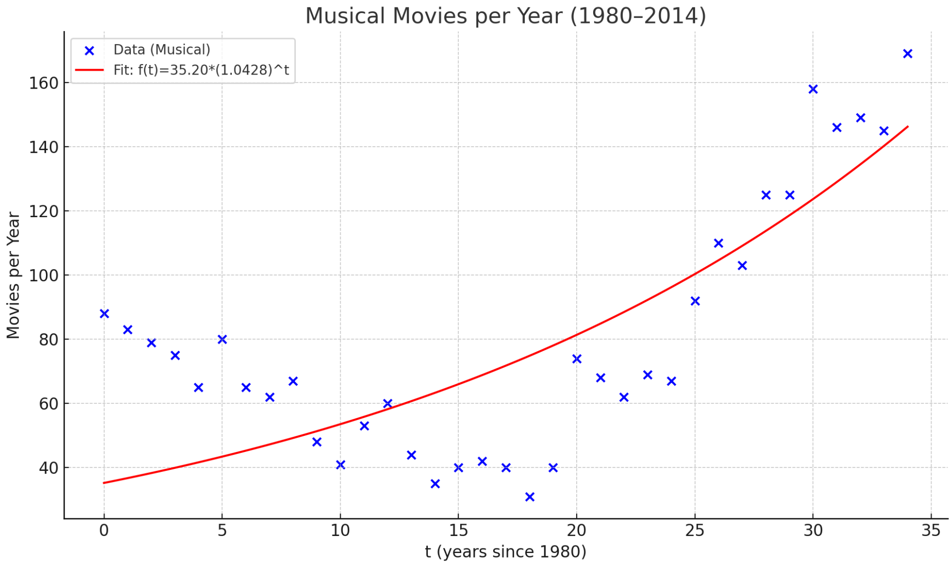
Use an LLM
Ask the LLM to generate Python code that plots the exponential models for your genre and the romance genre on the same axes.

4. According to the exponential models, will the annual number of movies in your chosen genre exceed that of romance movies? If yes, when and why? Use the graphs to support your answer. You can ask the LLM to extend the horizontal axis a few years to explore further. Explain how this is related to the yearly percent growth for each genre and the values of b.

Part 4: Other genres (some exponential, some not).

The file `movies_by_genre_1980_2014-moreGenres.csv` includes several additional genres, along with those from `movies_by_genre_1980_2014.csv`.

1. Ask the LLM to find an exponential model to fit the data of each genre in this new file and plot each genre's data with the model (one graph per genre).
2. The exponential model you get for "Musical" is not very good (see graph).
   In fact, in the graph we see that the actual number of "Musical" movies seems to have dropped from about 90 new movies in 1980 to about 40 new movies around 1995, and then started to increase. Because of this, the "Musical" genre does not seem to have experienced exponential growth between 1980 and 2014.



Musical movies 1980–2014

Find, among the plots you created, another genre which does not show exponential growth, and explain what about the data tells you it is not exponential.

3. Among the new genres, "Sport", "Sci-Fi", "Mystery", and "Animation" do appear to grow exponentially.
   a) Plot one of these together with Romance on a single graph (data and model for both).
   b) Although Romance looks larger throughout, your chosen genre has a higher growth rate. If the exponential trends continue, the number of movies each year in your genre will surpass the number of Romance movies! Estimate when this will happen using the graph (tell the LLM to extend the horizontal axis by four decades[2]) or by solving algebraically.

Further discussion

Make the math visible.

Ask the LLM to walk you in all detail through the math it did figure out the percent growth rate out of the value of b, and to build the sentence that says something like "4.736% growth per year" for Romance. Then ask it to show where that math is used in the code. Make sure you tell it you are a precalculus student and want to learn, so that it does not use any fancy math.

Provide an explanation in your own words.

All exponential genres for 1980–2014 in one plot.

1. Ask the LLM to plot all the exponential models from the genres that show exponential growth in the file `movies_by_genre_1980_2014-moreGenres.csv` on the same graph. Ask the LLM to include labels next to each curve to make them easy to identify.
2. From your graph, which genre seems to be growing the fastest? Double check with the growth rates to see if you are right—if not, in a couple of decades some other genre will overtake it! Ask the LLM to extend[3] the horizontal axis a couple of decades to see this.

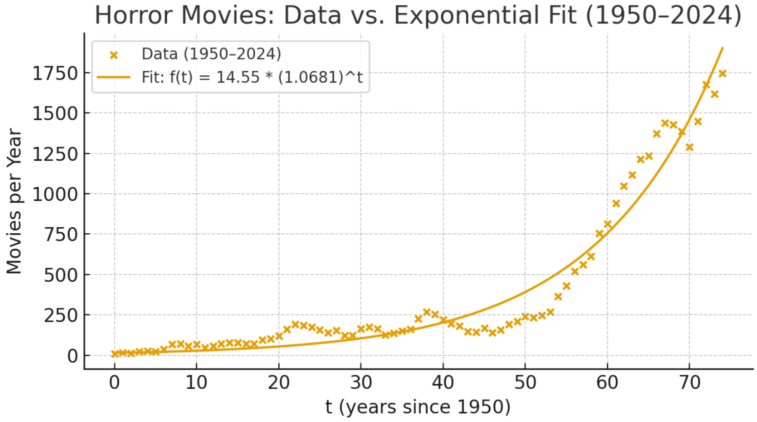
Genres for the full range 1950–2024 which are still growing exponentially.

The file `movies_by_genre_1950_2024.csv` includes data from 1950—2024.

1. Ask the LLM to plot the data from 1980 (not yet from 1950) to 2024 together with the exponential models you already created for 1980–2014 (one graph per genre).
2. From your graphs, do any of the models do a particularly bad job at predicting what happened in the decade 2014–2024? Select one and explain.

[2]This extension may be unreasonable—here we are focusing on exponential functions and how they relate to each other.
[3]Once again, we know this extension may be an unreasonable projection into the future.

3. Look for the impact of COVID-19 on movie production around 2020. What do you see? [Hint: what is the value of t for 2020?]
4. Although your graphs in part (b) suggest that most exponential models fail to give good predictions for the decade 2014—2024, in many cases the apparent slowdown during this period is simply part of variation that balances out over a longer time span (similar dips have occurred before). For instance, in the "Horror" genre, the slowdown in the decade 2014—2024 mirrors several previous dips, each of which was followed by a recovery. See the graph.



a) Ask the LLM to find a new exponential model to fit the data for each genre across the whole range 1950–2024 and plot data and model (one plot per genre)

b) Look for another genre, like "Horror", which seems to exhibit exponential growth over this entire period.

Sample Python Code

```
# --- Setup: imports ---
import numpy as np
import pandas as pd

# Plotly for interactive charts
import plotly.graph_objects as go

# Nonlinear least squares (fit exponential function to the data)
from scipy.optimize import curve_fit

# --- Load the CSV ---
FILENAME = "movies_by_genre_1980_2014.csv"

# In Colab, this will let you upload the file if it's not already present.
try:
    df = pd.read_csv(FILENAME)
except FileNotFoundError:
    try:
        # If you're in Google Colab, run the upload prompt
        from google.colab import files
        print(f"Couldn't find {FILENAME}. Please upload it in the dialog that opens.")
        uploaded = files.upload() # choose the CSV file from your computer
        df = pd.read_csv(FILENAME)
    except Exception as e:
        raise e

# --- Prepare the data ---
# Columns are capitalized in the file, e.g. "Romance", "Sci-Fi", etc.
# We'll extract Romance counts and use t = Year - 1980.
year = df["Year"].to_numpy()
romance = df["Romance"].to_numpy()

# Filter to the requested range 1980--2014 (inclusive), in case extra rows exist
mask = (year >= 1980) & (year <= 2014)
year = year[mask]
romance = romance[mask]
```

```python
# Independent variable t in years, with t=0 at 1980
t = year - 1980

# --- Define the model f(t) = a * b^t ---
def expo_model(t, a, b):
    # a > 0, b > 0; curve_fit will estimate best a, b
    return a * (b ** t)

# --- Initial guesses for (a, b) ---
# a: start near the first data point
a0 = max(1e-6, float(romance[0]))
# b: growth factor from first to last point (avoid zero divide)
if len(t) > 1 and romance[0] > 0:
    b0 = (romance[-1] / romance[0]) ** (1.0 / (t[-1] - t[0]))
    # Guard against degenerate values
    if not np.isfinite(b0) or b0 <= 0:
        b0 = 1.0
else:
    b0 = 1.0

# --- Fit the model directly (no logs) ---
# Constrain a > 0 and b > 0 for a proper exponential
popt, pcov = curve_fit(expo_model, t, romance, p0=[a0, b0], bounds=([1e-8, 1e-8], [np.inf, np.inf]))
a_hat, b_hat = popt

# --- Print the model in friendly form ---
print("Exponential model (t in years since 1980):")
print(f"f(t) = a * b^t")
print(f"a = {a_hat:.4f}, b = {b_hat:.6f}")
print(f"f(t) = {a_hat:.4f} * ({b_hat:.6f})^t")

# --- Build a smooth curve for plotting alongside the data ---
t_smooth = np.linspace(t.min(), t.max(), 400)
f_smooth = expo_model(t_smooth, a_hat, b_hat)

# --- Plot with Plotly (t on the horizontal axis) ---
fig = go.Figure()

# Data points
fig.add_trace(go.Scatter(
    x=t, y=romance,
    mode="markers",
    name="Romance movies (data)",
    hovertemplate="t=%{x}<br>count=%{y}<extra></extra>"
))

# Fitted curve
fig.add_trace(go.Scatter(
    x=t_smooth, y=f_smooth,
    mode="lines",
    name="Fitted exponential f(t)",
    hovertemplate="t=%{x:.2f}<br>f(t)=%{y:.2f}<extra></extra>"
))

fig.update_layout(
    title="Romance Movies per Year vs. Exponential Model (t = years since 1980)",
    xaxis_title="t (years since 1980)",
    yaxis_title="Number of Romance Movies",
    template="plotly_white",
    legend=dict(
        x=0.01,
        y=0.99,
        bgcolor="rgba(255, 255, 255, 0.5)",
        bordercolor="rgba(0, 0, 0, 1.0)", # Set border color to black
        borderwidth=1 # Set border width
    )
)

fig.show()
```