

# Inflation

November 17, 2021

Because the dollar amounts included in the IMDB Movie dataset's `worldwide_gross_income` column were not adjusted for inflation, it was necessary to make the adjustment in order to regularize the value of the gross income of the film so that meaningful predictions could be made. To do this, this project made use of inflation data collected by the U.S. Bureau of Labor Statistics from 1913-present.

A common measure of inflation over time is the Consumer Price Index. The Consumer Price Index is defined to be

$$CPI_t = \frac{C_t}{C_0} \cdot 100 \quad (1)$$

where  $C_t$  is the cost of a “basket” of goods in year  $t$  and  $C_0$  is the cost of a “basket” (of the same goods) in some base year. Thus, the ratio of the cost of an item in one year to the cost of the item in another year can be calculated using the ratio of the consumer price indices.

$$\begin{aligned} \frac{CPI_1}{CPI_2} &= \frac{C_1 \cdot 100}{C_0} \cdot \frac{C_0}{C_2 \cdot 100} \\ &= \frac{C_1}{C_2} \end{aligned}$$

Thus, a formula for the value of an amount after inflation is given by

$$C_2 = C_1 \frac{CPI_2}{CPI_1} \quad (2)$$

This project used the Bureau of Labor Statistics' Consumer Price Index dataset [1]. The dataset was loaded and then inner-joined to the IMDB dataset by `year` and the `Annual` column was selected to give the average CPI for the year that the film was released. Then, the CPI for 2020 was taken from the dataset and equation (2) was used to calculate the gross income for the film, adjusted to the equivalent value for 2020.

A nearly identical process was followed for the `budget` column of the IMDB Movie dataset, however, before the adjustment for inflation could be applied, the currency amount needed to be converted to U.S. dollars.

## References

- [1] United States Bureau of Labor Statistics. *All items in U.S. city average, all urban consumers, not seasonally adjusted. Series ID CUUR0000SA0.* United States Bureau of Labor Statistics, 2020.