## Tackling Debt

1. **Based on your initial review of the data, does Mia pay off her credit card debt? Do you have any concerns about whether or not that is accurate?**

Mia successfully manages to pay off her debt using both methods, and I have no concerns about their accuracy.

1. **Does Mia reach the $10,000 limit set by her parents? Does the payment amount reflect the updated value of $2,000?**

Mia does not reach that $10,000 limit with either method.

1. **What chart types did you use in your EDA? Why did you choose those types?**

I utilized a line graph for the EDA because it effectively illustrates the trend of decreasing balances with payments over time.

1. **Was there any data that needed to be cleaned? What was it and why would it be considered “dirty” data?**

Most of the data was missing from the debt repayment plan charts. Additionally, the existing data was inaccurately calculated because the APR does not compound monthly at the full interest rate. APR represents the annual rate, so monthly it should be calculated as 20%/12 months = 1.67%.

1. **Were there any observations made during the EDA that helped you locate some dirty data?**

During the analysis, it became evident that the monthly amounts paid did not align with the given instructions, and there were inconsistencies in how interest was computed.

## Housing Prices

For the three factors you chose, answer the following three questions:

1. **Were there any data points that you needed to look up the historical context to understand?**

The CPI/Time deviation occurred due to the housing bubble in 2008, specifically the Sub-Prime Mortgage Crisis.

1. **What were some observations you made after calculating summary statistics?**

When calculating the average unemployment rate, it's important to exclude the upper quartile to address outliers, such as those seen during events like the onset of the pandemic in 2020.

1. **Was there any “dirty” data you found in the dataset that you would need to handle?**

Some data points for unemployment rates are represented as decimals, which can lead to skewed results during summarization (e.g., 3.4 could be misinterpreted as 340% instead of 3.4%).