**Hypothesis Testing Summary**

**Contributor: R. Holder**

**Objective:** The objective of this analysis was to determine if there are significant differences in the mean values of satellite positioning data (gse\_x\_ace, gse\_y\_ace, gse\_z\_ace) across different periods (train\_a, train\_b, etc.) using ANOVA (Analysis of Variance) tests.

**Steps Taken:**

1. **Data Preparation:**
   * Loaded the dataset containing satellite positioning data from a CSV file.
   * Filtered data for each period (train\_a, train\_b, etc.) to create separate datasets (period\_a\_data, period\_b\_data, etc.).
2. **Hypothesis Testing:**
   * Conducted ANOVA tests using the f\_oneway function from scipy.stats for each satellite positioning column (gse\_x\_ace, gse\_y\_ace, gse\_z\_ace).
   * Calculated the F-statistic and p-value for each test.
3. **Results:**
   * **gse\_x\_ace:**
     + F-statistic: 1.47
     + p-value: 0.226
     + Interpretation: The p-value (0.226) suggests that there is no significant difference in the mean gse\_x\_ace values across the periods (train\_a, train\_b, etc.).
   * **gse\_y\_ace:**
     + F-statistic: 0.89
     + p-value: 0.344
     + Interpretation: The p-value (0.344) indicates no significant difference in the mean gse\_y\_ace values across the periods.
   * **gse\_z\_ace:**
     + F-statistic: 0.77
     + p-value: 0.382
     + Interpretation: The p-value (0.382) shows no significant difference in the mean gse\_z\_ace values across the periods.

**Conclusion:** Based on the ANOVA tests conducted, there is no statistically significant evidence to suggest that the mean values of gse\_x\_ace, gse\_y\_ace, or gse\_z\_ace differ significantly across the different periods (train\_a, train\_b, etc.). Therefore, we fail to reject the null hypothesis for all tested variables.

**Next Steps:**

* Further exploration could involve additional statistical tests or deeper analysis to investigate other potential factors influencing satellite positioning data.
* Consider exploring correlations or trends within each period or across different datasets.
* Visualize the data to gain more insights and communicate findings effectively.

This analysis provides a foundational understanding of the variability in satellite positioning data across different time periods specified in the dataset.