**Project: Satellite Positioning Analysis**

**Contributor: R.Holder – Data Scientist**

**Objective:** Create an analysis project to explore and analyze satellite positioning data, specifically focusing on understanding the variations in satellite positions over time.

**Steps:**

**Step 1: Data Acquisition and Preparation**

* **Objective:** Obtain satellite positioning data for analysis.
* **Actions Taken:**
  + Acquired the dataset from a specified file path (C:/Users/rosek/Desktop/Satellite Positioning.csv).
  + Imported necessary libraries (pandas, matplotlib, math).
  + Loaded the dataset into a Pandas DataFrame (df) for initial exploration.
  + Checked basic information about the dataset using df.head(), df.info(), and df.describe() to understand its structure, column names, data types, and missing values.

**Step 2: Data Cleaning and Transformation**

* **Objective:** Prepare the dataset for analysis by handling missing data and converting necessary columns.
* **Actions Taken:**
  + Converted the timedelta column to timedelta64 format to facilitate time-based analysis.
  + Conducted initial data cleaning to handle missing values, especially in columns such as gse\_x\_dscovr, gse\_y\_dscovr, and gse\_z\_dscovr.

**Step 3: Data Exploration and Visualization**

* **Objective:** Explore the dataset visually and statistically to understand trends and patterns in satellite positioning.
* **Actions Taken:**
  + Plotted the variation of gse\_x\_ace over time (df['timedelta'] vs df['gse\_x\_ace']) using matplotlib.
  + Analyzed descriptive statistics (df.describe()) for numerical columns to understand the range and distribution of data.

**Step 4: Hypothesis Testing**

* **Objective:** Test hypotheses regarding differences in satellite positions across different periods.
* **Actions Taken:**
  + Filtered the dataset to create subsets (period\_a\_data, period\_b\_data, etc.) based on the period column.
  + Conducted ANOVA tests using scipy.stats.f\_oneway to determine if there are significant differences in mean values of gse\_x\_ace, gse\_y\_ace, and gse\_z\_ace across different periods.
  + Interpreted results based on calculated F-statistics and p-values to understand statistical significance.

**Step 5: Conclusion and Next Steps**

* **Objective:** Summarize findings and outline next steps for further analysis.
* **Actions Taken:**
  + Concluded that there were no significant differences in the mean values of gse\_x\_ace, gse\_y\_ace, and gse\_z\_ace across the periods (train\_a, train\_b, etc.).
  + Recommended further exploration, such as investigating correlations within each period or exploring additional datasets for comparative analysis.
  + Discussed potential visualizations and statistical techniques for deeper insights into satellite positioning dynamics.

.