**Requirements Document for Marvel Character Dataset Analysis**

**Objective**

The primary goal of this project is to conduct a comprehensive analysis of a Marvel character dataset to uncover trends, relationships, and patterns related to character traits, popularity, survival, and their evolution over time.

**Scope of Work**

This analysis will encompass:

* Data cleaning and preprocessing
* Descriptive statistics and visualization
* Correlation analysis between character traits
* Time-based analysis of character trends
* Trait-based popularity and survival analysis
* Advanced insights (optional)

**Stakeholders**

* **Primary Stakeholders:** Marvel enthusiasts, comic book researchers, media analysts
* **Secondary Stakeholders:** Data scientists, analysts, machine learning enthusiasts

**Deliverables**

* **Data Cleaning and Preprocessing:** Handle missing values, ensure proper data types, remove or impute missing data.
* **Descriptive Analysis and Visualization:** Provide basic statistics, visualize key character traits, generate charts.
* **Correlation and Comparative Analysis:** Analyze correlations between character traits, investigate relationships.
* **Time-based Analysis of Character Evolution:** Plot character debuts per year and decade, explore trait evolution over time.
* **Popularity and Survival Analysis:** Identify top characters, explore survival trends across character types.
* **Advanced Insights (Optional):** Investigate relationships using correlation heatmaps, build predictive models, explore network graphs.

**Functional Requirements**

* **Data Cleaning and Preprocessing:** Handle missing values, ensure data types, filter unnecessary columns.
* **Descriptive Statistics and Visualization:** Generate summary statistics, create visualizations.
* **Correlation and Comparative Analysis:** Calculate correlations, compare traits, explore trends.
* **Time-based Analysis of Character Evolution:** Filter dataset by year, visualize character debuts, track changes.
* **Popularity and Survival Analysis:** Sort characters, analyze survival status, compare across alignments.
* **Advanced Insights (Optional):** Apply machine learning models, build network graphs.

**Non-Functional Requirements**

* **Performance:** Handle large datasets efficiently, ensure quick visualization loading.
* **Usability:** Provide intuitive visualizations, deliver actionable insights.
* **Scalability:** Be extendable to handle additional datasets.

**Milestones & Timeline**

* **Week 1:** Data cleaning and preprocessing, initial summary statistics and basic visualizations.
* **Week 2:** Trait-based analysis, correlation analysis.
* **Week 3:** Time-based analysis, survival and popularity analysis.
* **Week 4:** Advanced insights and predictive modeling (optional), final report and presentation.

**Technology Stack**

* **Languages/Tools:** Python
* **Libraries:** Pandas, Seaborn, Matplotlib
* **IDE/Platform:** Jupyter Notebooks or PyCharm

**Acceptance Criteria**

* All missing data handled appropriately.
* Descriptive statistics and visualizations provided for all key character traits.
* Correlations and relationships between traits clearly analyzed and visualized.
* Time-based trends visualized effectively.
* Optional: Predictive models implemented with acceptable accuracy.

**This document outlines the complete analysis framework, ensuring that each step of the process delivers actionable insights from the Marvel character dataset.**