

## **My Data Processing and Analysis of Spell Pricing in Final Fantasy 1**

### **Objectives**

My goal with this project is to analyze and process spell data from Final Fantasy 1 to predict spell prices based on key attributes. The process involves feature selection, exploratory data analysis (EDA), data transformation, outlier detection, and model training.

### **Data Loading and Feature Selection**

I started by loading the dataset (`ff1\_spells.csv`), which contains various attributes of spells. However, for this study, I decided to keep only the following features:

- level
- mp
- power
- accuracy
- target
- price (dependent variable)

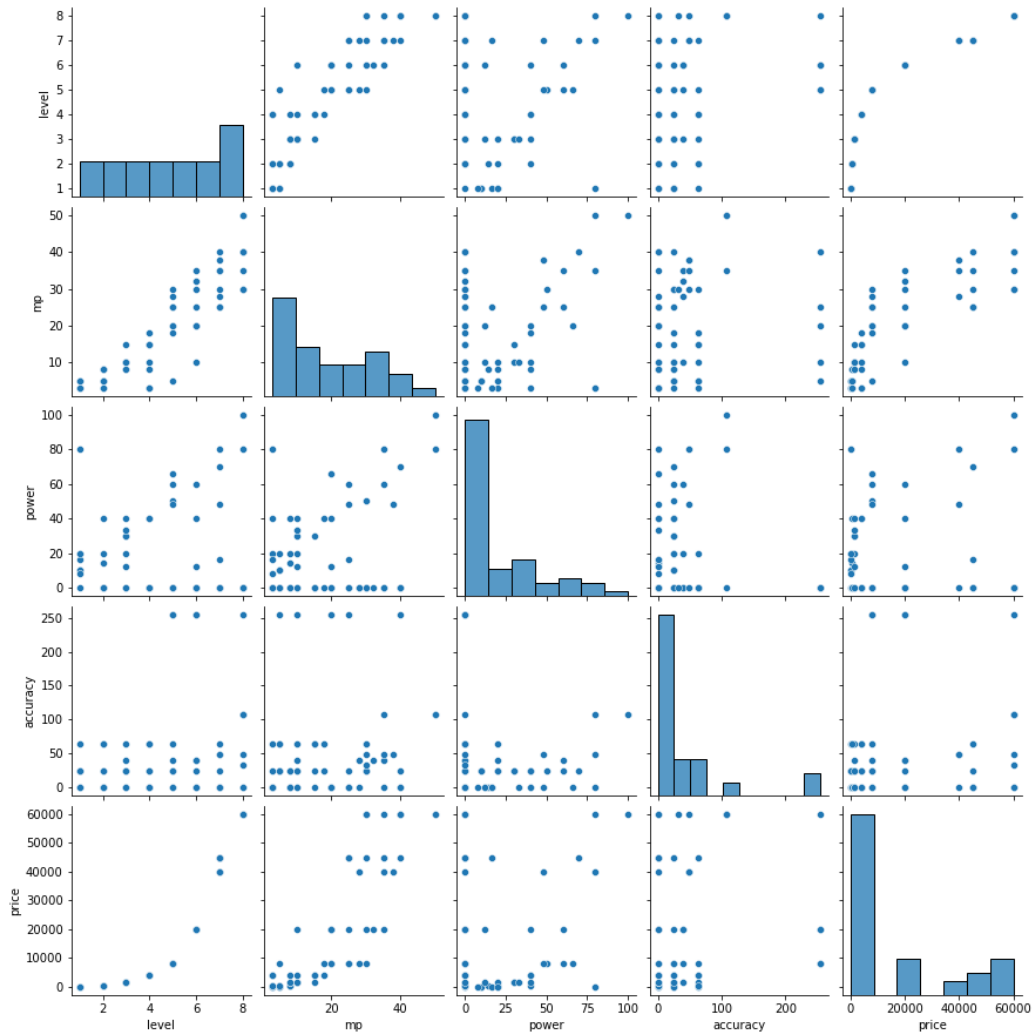
All other features were removed. I then inspected the data for missing values and unique values to better understand its structure.

### **Exploratory Data Analysis (EDA)**

To better understand the dataset, I conducted various analyses.

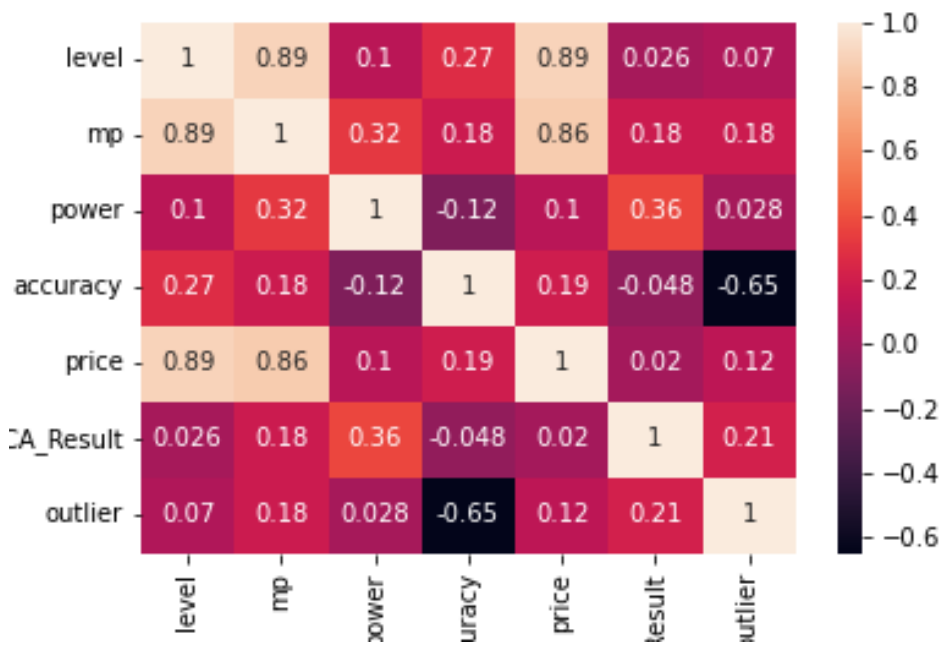
### Histogram of Price Feature

I plotted a histogram of the price feature to visualize its distribution. This helped me identify the spread of price values, presence of skewness, and potential outliers.



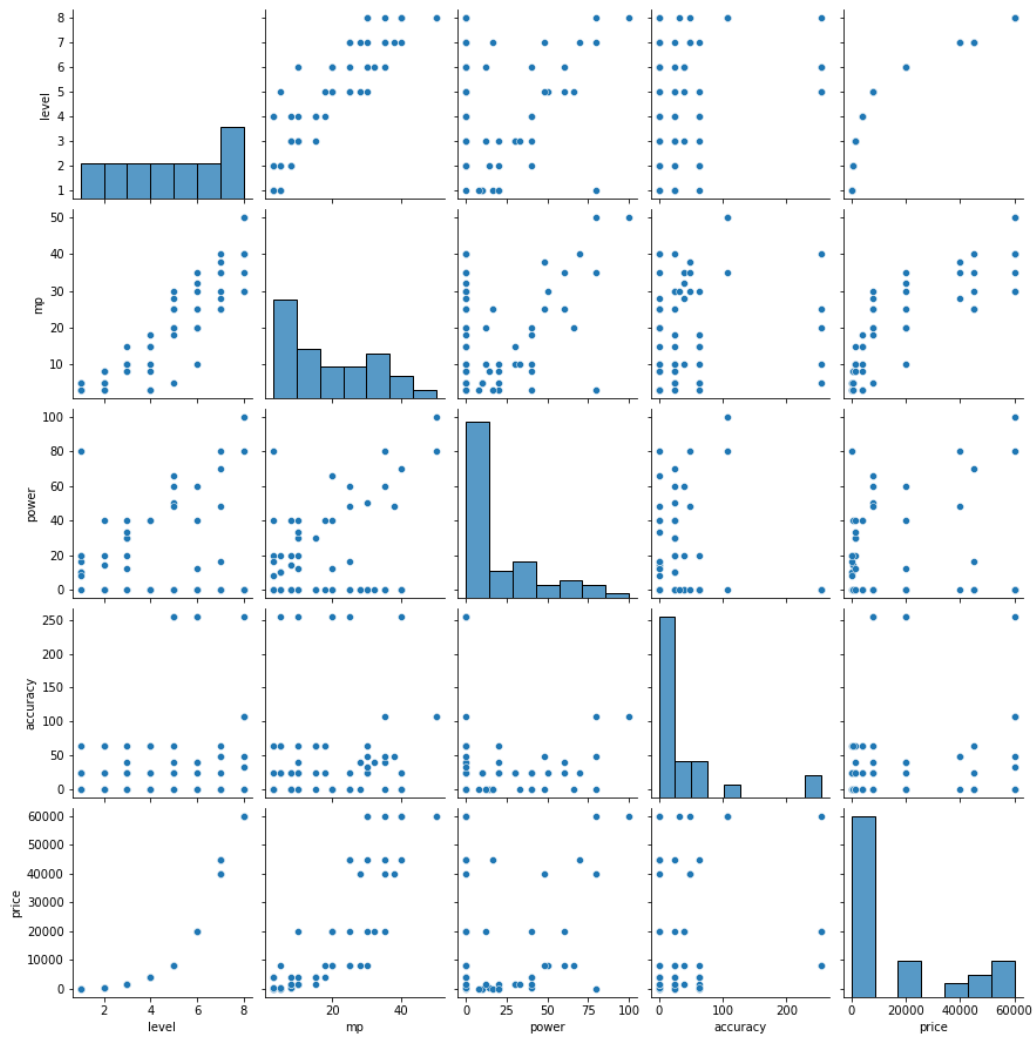
### Correlation Plot

Next, I generated a correlation plot to examine relationships between different numerical variables. This was useful in determining which features had strong correlations with the target variable (price).



## Pair Plot

I also created a pair plot to observe relationships between features and their distributions.



## Data Processing

### Encoding Categorical Features

Since target is a categorical feature, I converted it into a one-hot encoded array. To avoid excessive dimensionality, I applied Principal Component Analysis (PCA) to reduce the one-hot encoded matrix into a single-dimensional numerical feature. I then added this transformed feature to the dataset and removed the original target column.

### Outlier Detection

To detect outliers, I used the Local Outlier Factor (LOF) method. Each data point was assigned an outlier score, helping me identify anomalies that could impact model training.

### Feature Selection

To determine the most relevant features for predicting price, I used the SelectKBest method. This ranked the features based on their correlation with the target variable. The top two features were identified, but I chose to keep the dataset unchanged at this stage.

## Model Training

### Dummy Regressor

For my baseline model, I used a Dummy Regressor. The model was trained using level and mp as input features, with price as the output. After evaluating the model's performance, I made a test prediction using the following input values:

- [1, 5, 20, 50, 0.5]

The predicted price from this model serves as a reference point for further improvements.

## Conclusion

This study outlines my structured approach to processing and analyzing spell pricing in Final Fantasy 1. By conducting exploratory data analysis, encoding categorical variables, detecting outliers, selecting the best features, and training a baseline model, I established a foundation for improving model accuracy and prediction reliability.