

```
python
```

```
import pandas as pd
```

```
def load_data(filepath):
```

```
    return pd.read_csv(filepath)
```

```
def handle_missing_values(data):
```

```
    # Fill numeric columns with the mean
```

```
    numeric_cols = data.select_dtypes(include=['number']).columns
```

```
    for col in numeric_cols:
```

```
        data[col].fillna(data[col].mean(), inplace=True)
```

```
    # Fill categorical columns with the mode
```

```
    categorical_cols = data.select_dtypes(include=['object']).columns
```

```
    for col in categorical_cols:
```

```
        data[col].fillna(data[col].mode()[0], inplace=True)
```

```
def remove_duplicates(data):
```

```
    initial_count = data.shape[0]
```

```
    data.drop_duplicates(inplace=True)
```

```
    print(f"Removed {initial_count - data.shape[0]} duplicates")
```

```
def correct_data_formats(data):
```

```
    if 'Date' in data.columns:
```

```
        data['Date'] = pd.to_datetime(data['Date'])
```

```
    if 'Category' in data.columns:
```

```
        data['Category'] = data['Category'].astype('category')
```

```
def save_clean_data(data, path):
```

```
    data.to_csv(path, index=False)
```

```
    print("Cleaned data saved successfully.")
```

Example usage

```
if __name__ == "__main__":
```

```
    data = load_data('path/to/your/dataset.csv')
```

```
    handle_missing_values(data)
```

```
    remove_duplicates(data)
```

```
    correct_data_formats(data)
```

```
    save_clean_data(data, 'path/to/cleaned_data.csv')
```