

## Appendix II: Numerical explanatory variables plotted against each other

```
In [1]: """Imports necessary packages"""

import itertools
import math
from typing import Dict, Iterable, List, Union

import matplotlib.pyplot as plt
import numpy as np
import pandas as pd
import pylab
import scipy
import scipy.stats as stats
import seaborn as sns
import statsmodels.api as sm
from sklearn.model_selection import train_test_split

sns.set_style("whitegrid")
```

```
In [2]: def make_scatterplot(x_data: Iterable, y_data: Iterable, xlabel: str = "x", ylabel:
        """Prints a scatterplot.

        Args:
            x_data (Iterable): the one dimensional data to plot on the x axis.
            y_data (Iterable): the one dimensional data to plot on the y axis.
            xlabel (str, optional): the label for the x axis. Defaults to "x".
            ylabel (str, optional): the label for the y axis. Defaults to "y".
            title (str, optional): the title of the plot. Defaults to "A scatterplot".
        """
        plt.figure()
        plt.scatter(x_data, y_data)
        plt.xlabel(xlabel)
        plt.ylabel(ylabel)
        plt.title(title)
        plt.show()
```

```
In [3]: data = pd.read_csv("D:/School/frequentist-statistics/ITM-song-popularity/database/it
data = data.drop("Unnamed: 0", axis=1)
```

```
In [4]: numerical_variables = ["name_len", "track_number", "duration", "acousticness", "danc
categorical_variables = ["key", "mode", "time_signature", "explicit"]
```

```
In [5]: num_combs = itertools.combinations(numerical_variables, 2)
for x, y in num_combs:
    make_scatterplot(data[x], data[y], xlabel=x, ylabel=y, title="%s vs %s" % (x, y))
```





















































































