1. How should one approach the Travel Package Purchase Prediction project?

- Before starting the project, please read the problem statement carefully and go through the criteria and descriptions mentioned in the rubric.
- Once you understand the task, download and import the dataset into a Jupyter notebook to get started with the project.
- To work on the project, you should start with an exploratory analysis of the data. Try to use descriptive statistics and visualization to understand the data.
- Once the EDA is completed, you should start building possible models (tuned and untuned) and test for corresponding performance metrics to find the best-performing model.
- It is important to close the analysis with key findings and recommendations to the business.

2. Is the "Fe male" entry the same as "Female" in the Gender column? If yes, how to treat them?

Yes, the entries "Fe male" and "Female" are the same. Considering it as a data entry error, it is recommended to convert "Fe male" into "Female".

3. I am getting the below error with the Decision Tree classifier:

```
ValueError: could not convert string to float: 'Self Enquiry'
```

How to solve it?

As decision trees cannot handle non-numeric data, it is preferable to either encode or create dummy variables for the categorical features.

4. How to install XGBoost on Windows?

To install XGBoost on Windows, use "!pip install xgboost" in Jupyter notebook or "conda install -c anaconda py-xgboost" in the Anaconda prompt in Anaconda.

Before this, it must be ensured that previous versions of XGBoost are uninstalled by using the command "!pip uninstall xgboost".

5. How to install XGBoost on Mac?

To install XGBoost on Mac, run the following command in terminal "conda install -c conda-forge xgboost".

In some cases, trying to install XGBoost may return the following error:

```
XGBoostError: XGBoost Library (libxgboost.dylib) could not be loaded.
Likely causes:
  * OpenMP runtime is not installed (vcomp140.dll or libgomp-1.dll for Windows, libomp.dylib for Mac OSX, libgomp.so for Linux and
  * You are running 32-bit Python on a 64-bit OS
```

In this scenario, run the following commands on the Terminal on macOS:

Note: Open a new Terminal window before using the below commands in the macOS Terminal.

To open the Terminal, open the Launchpad and then click on the Terminal icon.

1. Install the Homebrew package manager by running the following command on the Terminal.

```
/bin/bash -c "$(curl -fsSL https://raw.githubusercontent.com/Homebrew/install/HEAD/install.sh)"
```

2. After this, use Homebrew to install the OpenMP library by running the following command on the Terminal.

```
brew install libomp
```

This should allow you to install XGBoost on your Mac.

6. Are "unmarried" and "single" in "Marital status" the same? If not how?

Unmarried and single are not the same in the marital status feature, since an "unmarried" person could potentially have a partner but not be married to them, whereas a "single" person is merely one without a partner.

7. What is the currency of 'MonthlyIncome'? Is it too much to be in US dollars?

As there is no currency information mentioned in the data, you are free to use a currency of your choice. This does not make a difference to the algorithm's predictive capabilities.

8. Error in reading CSV file Tourism because the file extension is .xlsx. Any solution?

The file is in .xlsx format (Microsoft Excel file) and not .csv. You need to write the extension of the file as .xlsx and you can then use the function **pd.read_excel()** to import the file.

9. Tuning XGBoost is taking too much time. What to do?

Please find below the options that can be tried to reduce the run time for XGBoost tuning:

1. Pick individual features at a time and try adjusting them. Get the optimal value beyond which no improvement is seen, note that optimal value, and move on to the next feature.
2. Only use parameters that are necessary. Do not use all the parameters just because they are there.
3. Do not use more than 2-3 options in each list of hyperparameters given to the tuning function.
4. Try a small dataset randomly chosen from the original dataset, tune hyperparameters on them, and then use that set of hyperparameters over the entire dataset.

10. I am getting the below error with Bagging Classifier with Logistic Regression as base_estimator:

```
AttributeError: 'str' object has no attribute 'decode'
```

How to solve it?

The LogisticRegression function has a parameter named solver, which was earlier by default set to 'liblinear', but after an update in sklearn, the default solver was changed to 'lbfgs'. Kindly try setting the solver as liblinear, and try the code again, as shown below:

```
bagging_lr=BaggingClassifier(base_estimator=LogisticRegression(solver='liblinear', random_state=1),random_state=1)
bagging_lr.fit(X_train,y_train)
```

11. I am getting the below warning while fitting XGBoost Classifier

```
WARNING: Starting in XGBoost 1.3.0, the default evaluation metric used with the objective 'binary:logistic' was changed from 'erro
```

How to solve it?

To remove the warning kindly try setting the eval_metric hyperparameter as 'logloss', as shown below:

```
xgb = XGBClassifier(eval_metric='logloss')
```