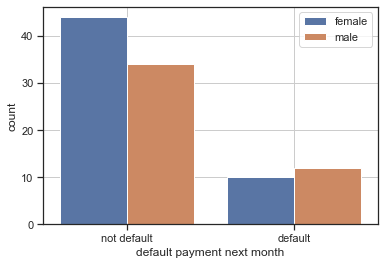
**Credit Data EDA Report**

Below are few graphs that show the relationship between different variables and their meanings.

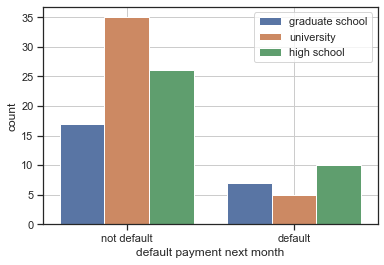
**Default payment** **vs. Gender**



*Not Default* - more females tend to pay on time than males

*Default* - more males tend to default payment than females

**Default Payment vs. Education**

****

*Not Default*

number of clients with different level of education from high to low

1. University

2. high school

3. graduate school

*Default*

number of clients with different level of education from high to low

1. high school

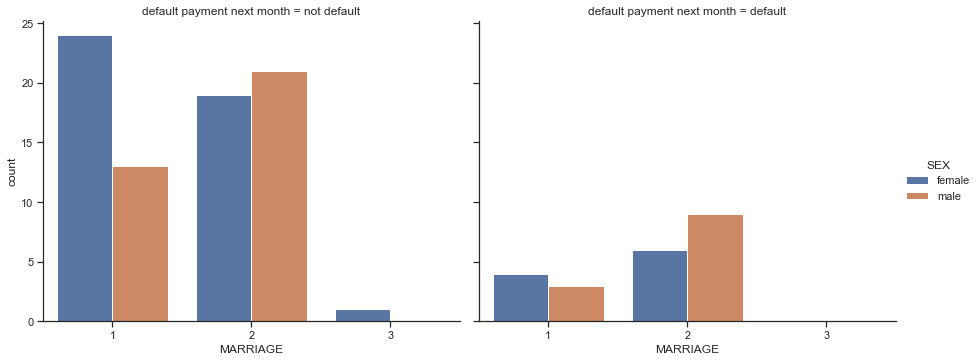
2. graduate school

3. university

To sum up, clients with university degree tend to default less, clients with high school level

education tend to default more.

**Default Payment vs. Marriage with Sex**

*Not Default*

- Married females tend not to default more than married males

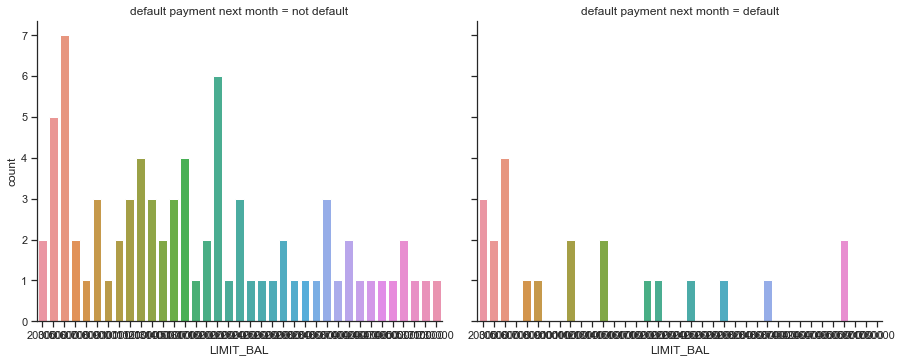
- Single males tend not to default more than females

*Default*

- Married females tend to default more than males

- Single males tend to default more than females

**Default Payment vs. Limit\_Balance**

****

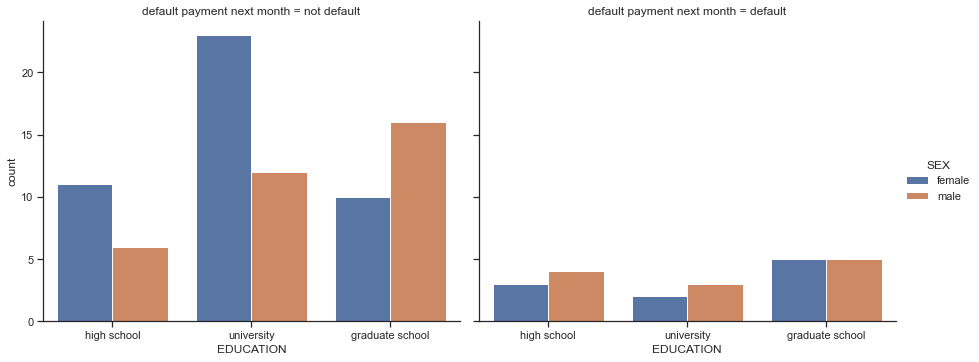
*Not Default*

Clients are given higher credit limit more frequently

Default

Clients are given lower credit limit less frequently

**Default Payment vs. Education with Sex**



*Not Default*

- Female clients with **University** degree tend not to default more than male clients

- Male clients with **Graduate school** education tend not to default more than female clients

- Female clients with **High school** education tend not to default more than males

*Default*

- Male clients with **University** degree tend to default more than females

- Number of clients with **Graduate** school education seems to be even with default payment

- Males clients with **High** school education tend to default more than females

**Correlation Relationship between Different Variables**

The Correlation relationship between different default payment and other variables such as Limit\_bal, Sex, Education etc. are very weak and even there is little to none relationship, which indicates that these factors probably wouldn’t affect whether customer pay their bills on time that much, which also means that when we apply predictive model to predict the result, the outcome maybe not be accurate.