Министерство образования Республики Беларусь

Учреждение образования

“Брестский государственный университет”

Кафедра ИИТ

Лабораторная работа №11

По дисциплине: “ЯП”

Тема: “Pandas”

Выполнил:

Студент 2-го курса

Группы ПО-7

Качан В.В.

Проверила:

Дряпко А.В.

Брест, 2021

Задание 1.

Для загрузки датасета pandas используем:

df = pd.read\_csv("marketing\_campaign.csv", delimiter="\t")  
df.sample(15)

Задание 2.

Для вывода общей информации о датасете используем:

df.info()

Задание 3.

Для проверки на NULL-данные используем:

df[df.isnull().T.any()]

Задание 4.

Для удаления колонки "Z\_CostContact", "Z\_Revenue" используем:

df.drop(["Z\_CostContact", "Z\_Revenue"], axis=1, inplace=True)

Задание 5.

Для переименования колонки "Year\_Birth" в "Age" используем:

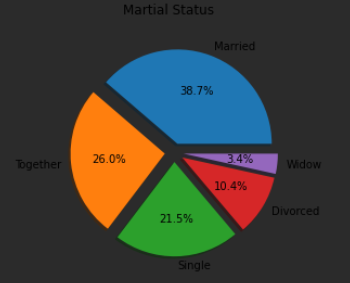
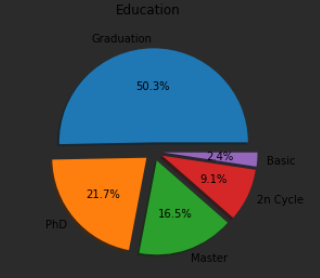
df.rename({"Year\_Birth": "Age"}, axis=1, inplace=True)

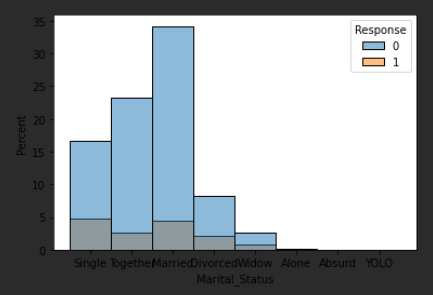
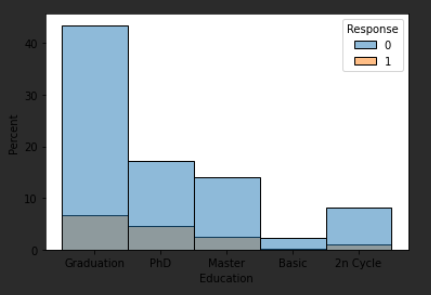
Задание 6.

Код программы:

df.drop(["Z\_CostContact", "Z\_Revenue"], axis=1, inplace=True)  
df.rename({"Year\_Birth": "Age"}, axis=1, inplace=True)  
df.info()  
  
ms\_df\_c = df["Marital\_Status"].value\_counts().to\_frame()  
ms\_df = ms\_df\_c[ms\_df\_c["Marital\_Status"] > 50]  
ed\_df = df["Education"].value\_counts().to\_frame()  
  
ms\_df[ms\_df["Marital\_Status"] > 50].plot(  
 kind="pie",  
 autopct="%1.1f%%",  
 legend=None,  
 ylabel="",  
 subplots=True,  
 shadow=True,  
 explode=(0.1, 0.1, 0.1, 0.1, 0.1),  
 title="Martial Status",  
)  
plt.show()  
  
ed\_df.plot(  
 kind="pie",  
 autopct="%1.1f%%",  
 legend=None,  
 ylabel="",  
 subplots=True,  
 shadow=True,  
 explode=(0.1, 0.1, 0.1, 0.1, 0.1),  
 title="Education",  
)  
plt.show()  
  
sns.histplot(data=df, x="Marital\_Status", stat="percent", hue="Response")  
plt.show()  
sns.histplot(data=df, x="Education", stat="percent", hue="Response")  
plt.show()

Результат выполнения:

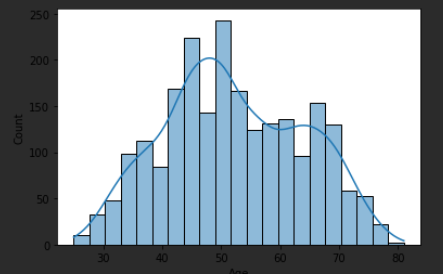
 

Задание 7.

Код программы:

def calculate\_age(born) -> int:  
 return int(datetime.today().strftime("%Y")) - int(born)  
  
df["Age"] = df["Age"].apply(calculate\_age)  
filtered\_df = df[df["Age"] < 100]  
sns.histplot(x=filtered\_df["Age"], kde=True)  
plt.show()

Результат выполнения:

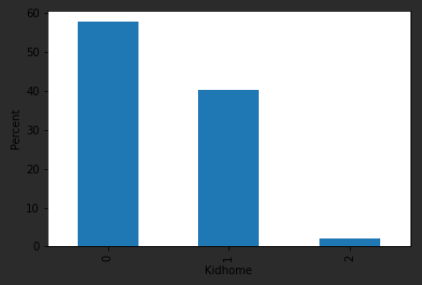
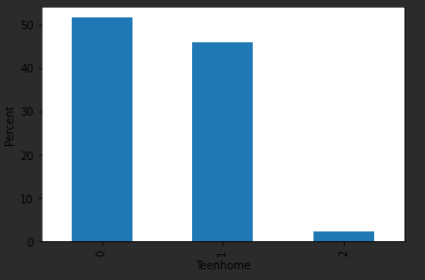


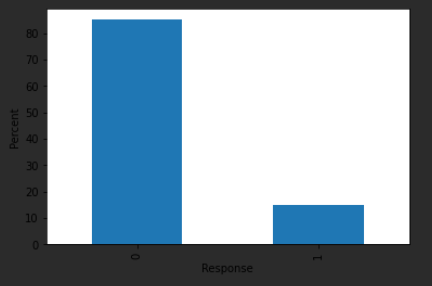
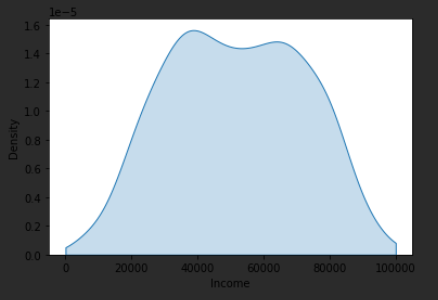
Задание 8.

Код программы:

def show\_hist(data\_name: str) -> None:  
 df[data\_name].value\_counts(normalize=True).mul(100).to\_frame().plot(  
 kind="bar",  
 legend=None,  
 xlabel=data\_name,  
 ylabel="Percent"  
 )  
 plt.show()  
  
show\_hist("Kidhome")  
show\_hist("Teenhome")  
show\_hist("Response")  
sns.kdeplot(df["Income"], shade=True, clip=(10\*\*0, 10\*\*5))  
plt.show()

Результат выполнения:

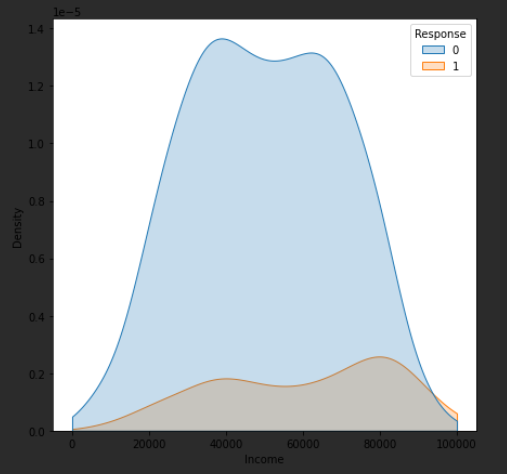
 

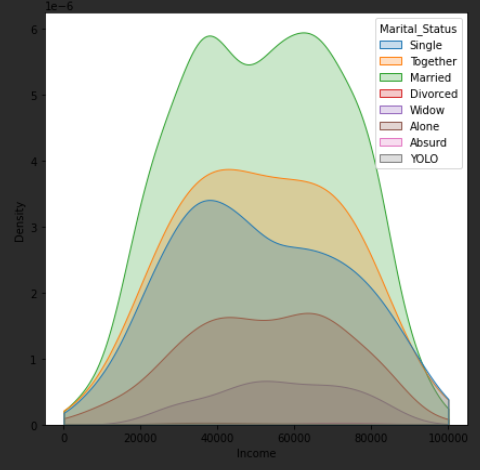
Задание 9.

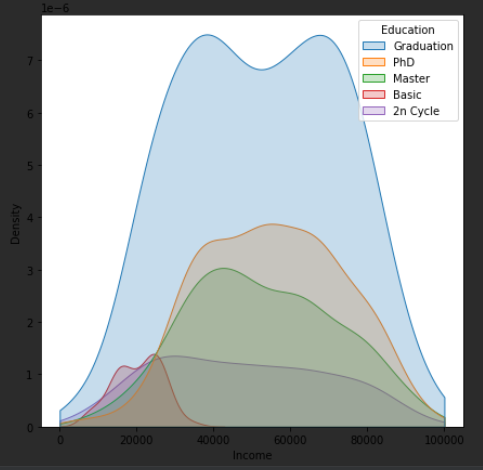
Код программы:

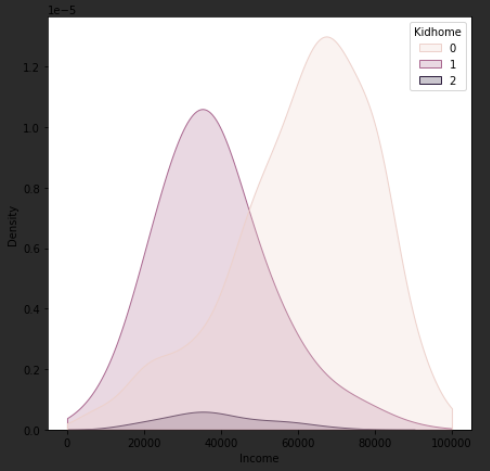
def show\_kdeplot(hue: str, clip: tuple = (10\*\*0, 10\*\*5)):  
 plt.figure(figsize = (7, 7))  
 sns.kdeplot(data=df, x="Income", hue=hue, shade=True, clip=clip)  
 plt.show()  
  
show\_kdeplot("Response")  
show\_kdeplot("Marital\_Status")  
show\_kdeplot("Education")  
show\_kdeplot("Kidhome")  
plt.show()

Результат выполнения:









Задание 10.

Код программы:

numerics = ["int16", "int32", "int64", "float16", "float32", "float64"]  
newdf = df.select\_dtypes(include=numerics)  
plt.figure(figsize=(20, 20))  
sns.heatmap(newdf.corr(), cmap="coolwarm", annot=True)  
plt.show()