#Program 1

Which of the variables have missing values?

```
import pandas as pd

# Read the CSV file into a DataFrame

df = pd.read_csv("People Charm case.csv")

# Check for missing values in each column

missing_values = df.isna().sum()

# Display columns with missing values

columns_with_missing = missing_values[missing_values > 0]

print(columns_with_missing)
```

#Program 2

#What is the third quartile value for the variable lastEvaluvation

```
import pandas as pd

# Read the CSV file into a DataFrame

df = pd.read_csv("People Charm case.csv")

# Calculate the third quartile for the "lastEvaluvation" variable

third_quartile = df["lastEvaluation"].quantile(0.75)

print("Third Quartile (75th Percentile) for 'lastEvaluvation':", third_quartile)
```

#Program 3

Construct a Crosstable for the variables dept and salary and find out which department has highest frequency value in the category low salary.

```
# Read the CSV file into a DataFrame

df = pd.read_csv("People Charm case.csv")

# Construct a crosstable for 'dept' and 'salary' variables

crosstab_result = pd.crosstab(df['dept'], df['salary'])

# Find the department with the highest frequency value in the 'low' salary category

department_with_highest_low_salary = crosstab_result['low'].idxmax()

print("Department with the highest frequency of low salary:", department_with_highest_low_salary)
```

#Program 4

#Generate a boxplot for the variable numberOfProjects and get the median value for the number of projects where the employees have worked on.

```
import pandas as pd
import matplotlib.pyplot as plt

# Read the CSV file into a DataFrame
df = pd.read_csv("People Charm case.csv")

# Generate a boxplot for the 'numberOfProjects' variable
plt.figure(figsize=(8, 6))
plt.boxplot(df['numberOfProjects'])
plt.title('Boxplot of numberOfProjects')
plt.ylabel('Number of Projects')
plt.show()

# Calculate the median value for the 'numberOfProjects' variable
median_number_of_projects = df['numberOfProjects'].median()

print("Median number of projects:", median_number_of_projects)
```

#Program 5

#Plot a histogram using the variable avgMonthlyHours and find the range in which the number of employees worked for 150 hours per month?

```
import pandas as pd
import matplotlib.pyplot as plt
import numpy as np

# Read the CSV file into a DataFrame
df = pd.read_csv("People Charm case.csv")
```

```
# Plot a histogram for the 'avgMonthlyHours' variable
plt.figure(figsize=(8, 6))
plt.hist(df['avgMonthlyHours'], bins=20, edgecolor='black')
plt.title('Histogram of avgMonthlyHours')
plt.xlabel('Average Monthly Hours')
plt.ylabel('Frequency')
plt.show()

# Find the range in which the number of employees worked for 150 hours per month
range_for_150_hours = df[(df['avgMonthlyHours'] >= 150) & (df['avgMonthlyHours'] < 151)]
number_of_employees = len(range_for_150_hours)

print("Number of employees who worked for 150 hours per month:", number_of_employees)

#Program 6
# Generate a boxplot for the variables lastEvaluation and numberOfProjects

import pandas as pd
import matplotlib.pyplot as plt

# Read the CSV file into a DataFrame
df = pd.read_csv("People Charm case.csv")

# Generate a boxplot for the 'lastEvaluation' and 'numberOfProjects' variables
plt.figure(figsize=(10, 6))
df.boxplot(column=['lastEvaluation', 'numberOfProjects'])
plt.title('Boxplot of lastEvaluation and numberOfProjects')
plt.ylabel('Values')
plt.show()
```