Here are the correct answers of Data Analyst MCQ Test:

1. What is a correct syntax to create a Pandas Series from a list?

   c) pd.Series(mylist)

2. Which of the following things can be data in Pandas?

   d) all of the mentioned

3. The result of an operation between unaligned Series will have the _____ of the indexes involved.

   b) union

4. Which of the following inputs can be accepted by DataFrame?

   d) All of the mentioned

5. Point out the wrong statement.

   b) Series can be passed into most NumPy methods expecting an ndarray (This statement is somewhat misleading; while Series can be used with many NumPy methods, there may be specific cases where it does not behave exactly like a ndarray).

6. Which of the following takes a dict of dicts or a dict of array-like sequences and returns a DataFrame?

   c) DataFrame.from_dict

7. If data is an ndarray, index must be the same length as data.

   a) True

8. What is a correct syntax to create a Pandas DataFrame?

   b) pd.DataFrame(data)

32. What is a correct syntax to add the labels "x", "y", and "z" to a Pandas Series?

b) pd.Series(mylist, index=["x", "y", "z"])

33. What is a correct syntax for changing the 'Firstname' column in the first row to 'John'?

a) df.loc[0, 'Firstname'] = 'John'

34. Consider the following code:

```
for x in df.index:
  if df.loc[x, 'Duration'] > 120:
    df.drop(x, inplace=True)
```

What will happen in the result when the value for 'Duration' is higher than 120?

c) The entire row will be deleted

35. What are duplicate rows in a DataFrame?

b) Identical rows

36. Insert the correct syntax for removing duplicates in a DataFrame.

df. drop_duplicates()

38. Insert the correct syntax for specifying that the plot should be of type 'scatter'.

df.plot(kind='scatter' , x='Duration', y='Calories')

39. Insert the correct syntax for specifying that the plot should be of type 'histogram'.

df['Duration'].plot(kind='hist')

41. True or False. If the number of rows is not specified, the tail() method will return all rows.

b) False (The tail() method returns the last 5 rows by default.)

42. What is a correct method for converting a column into date formats?

b) to_datetime()

43. True or false: A correlation of 0.9 is considered a good correlation.

a) True (0.9 indicates a strong positive correlation.)

44. Insert a correct syntax for finding relationships between columns in a DataFrame.

df. corr()

45. By default, if you print a large DataFrame with the print(df) function, Pandas will only return a selection of the data. How many rows will be returned?

a) 10, the headers and the first and last 5 rows

46. What is a correct syntax to add the labels "x", "y", and "z" to a Pandas Series?

b) pd.Series(mylist, index=["x", "y", "z"])

47. pandas was developed by?

c) Wes McKinney

48. What does the df.groupby() method do?

a) Groups rows based on column values for aggregation

49. What is the purpose of the df.fillna() method?

a) Fill missing values with specified data

50. How do you filter rows in a DataFrame based on a condition?

c) df[condition]

51. What does the df.pivot_table() method do?

a) Creates a pivot table for data aggregation

52. What does the df.agg() method do?

a) Aggregates data using one or more functions

53. How can you rename columns in a DataFrame?

a) df.rename(columns={'old_name': 'new_name'})

54. What does df.applymap(func) do?

a) Applies a function to each element of the DataFrame

55. How can you save a DataFrame to an Excel file?

a) df.to_excel('filename.xlsx')

56. What is the purpose of the df.duplicated() method?

a) Identifies duplicate rows in the DataFrame

57. How do you concatenate two DataFrames?

a) pd.concat([df1, df2])

58. What is the default color of the lines in a Matplotlib plot if no color is specified?

b) Blue

59. Which of the following is used to set the font size of labels in Matplotlib?

a) fontsize

60. How can you display multiple plots in a single figure in Matplotlib?

b) plt.subplots()

61. What does the plt.tight_layout() function do?

b) Automatically adjusts subplot parameters to give specified padding

62. Which Matplotlib function is used to add a grid to a plot?

a) plt.grid()

63. What will be output for the following code?

```
import pandas as pd
s = pd.Series([1,2,3,4,5], index=['a','b','c','d','e'])
print(s['a'])
```

a) 1

64. What will the following code do?

```
import pandas as pd
data = pd.read_csv('file.csv', usecols=['col1', 'col2'])
print(data.head())
```

b) Load only 'col1' and 'col2' from 'file.csv' and display the first five rows

65. What does this code return?

```
data.info()
```

b) Information about data types and non-null counts

66. What does this code snippet do?

```
import pandas as pd
data = pd.DataFrame({'A': [5, 10, 15], 'B': [1, 2, 3]})
print(data.describe())
```

a) Provides summary statistics for all numerical columns

67. What will be the result of the following code?

```
data['col1'] = data['col1'].apply(lambda x: x*2)
```

a) It doubles the values of 'col1' for all rows in the DataFrame

68. Which function is used to fill missing values with a specific value?

```
data['column'] = data['column'].____(value=0)
```

b) fillna

69. What does the following code do?

```
import matplotlib.pyplot as plt
data['column'].plot(kind='hist')
plt.show()
```

a) Displays a histogram of values in 'column'

70. What will be the output of the following code?

```
import numpy as np
arr = np.array([[1, 2], [3, 4], [5, 6]])
arr[1:, :1]
```

a) [[3], [5]]

71. What is the output of the following code?

```
import pandas as pd
data = pd.DataFrame({'A': [1, 2, 3], 'B': [4, 5, 6]})
data['C'] = data.apply(lambda x: x['A'] + x['B'], axis=1)
print(data['C'])
```

a) [5, 7, 9]

72. Given the following code, what will be the output of data.iloc[2] ?

```
import pandas as pd
data = pd.DataFrame({'X': [5, 6, 7], 'Y': [8, 9, 10]})
print(data.iloc[2])
```

b) X 7, Y 10

73. What will the following code return?

```
import numpy as np
arr = np.array([1, 2, 3, 4, 5])
result = arr[arr % 2 == 0]
```

b) [2, 4]

74. What will be the shape of the result after executing this code?

```
import numpy as np
arr = np.array([[10, 20, 30], [40, 50, 60]])
reshaped_arr = arr.reshape(3, 2)
print(reshaped_arr.shape)
```

b) (3, 2)

75. What will be the result of the following code?

```
import pandas as pd
df = pd.DataFrame({'A': [10, 20, 30], 'B': [5, 10, 15]})
df['C'] = df['A'] - df['B']
result = df[df['C'] > 10]
```

b) A DataFrame with one row

76. Given the following DataFrame, what will be the result of executing this code?

```
import pandas as pd
data = pd.DataFrame({'X': [3, 2, 1], 'Y': [9, 7, 8]})
result = data.sort_values(by='X')
print(result)
```

b) Data sorted by column 'X' in ascending order

77. What will be the result of the following code?

```
import pandas as pd
df = pd.DataFrame({'A': [1, 2, 3], 'B': [4, 5, 6]})
df['sum'] = df.apply(lambda x: x['A'] + x['B'], axis=0)
```

c) Error (You should use axis=1 for row-wise operation.)

78. What will be the result of the following merge operation?

```
import pandas as pd
left = pd.DataFrame({'ID': [1, 2, 3], 'Name': ['A', 'B', 'C']})
right = pd.DataFrame({'ID': [2, 3, 4], 'Score': [85, 90, 95]})
merged = pd.merge(left, right, on='ID', how='inner')
```

b) Keeps only rows with matching IDs in both DataFrames

79. What will be the result of the following code?

```
import numpy as np
arr = np.arange(10)
print(arr[::2])
```

b) [0, 2, 4, 6, 8]

80. What is the result of this slicing operation?

```
import numpy as np
```

```
arr = np.array([0, 1, 2, 3, 4, 5])
result = arr[1:5:2]
```

a) [1, 3]

81. What will be the result of the following pivot table operation?

```
import pandas as pd
data = pd.DataFrame({'A': ['foo', 'bar', 'foo'], 'B': ['one', 'two', 'one'], 'C': [1, 2, 3]})
pivot_table = pd.pivot_table(data, values='C', index='A', columns='B')
```

c) Both a and b (It aggregates and summarizes data based on unique combinations.)

82. What will be the output of the following code?

```
import numpy as np
arr = np.array([1, 2, 3, 4])
result = np.dot(arr, arr)
```

b) 30 (This is the dot product of the array with itself.)

83. Given the following data, what will be the result of the .loc[] operation?

```
import pandas as pd
data = pd.DataFrame({'X': [10, 20, 30], 'Y': [40, 50, 60]})
result = data.loc[0:1, 'Y']
```

b) [40, 50]

84. What does the following code do?

```
import pandas as pd
data = pd.DataFrame({'A': [1, 2, 3], 'B': [4, 5, 6]})
result = data[data['A'] > 1]
```

a) Filters rows where column 'A' has values greater than 1

85. What will be the output of this NumPy operation?

```
import numpy as np
arr = np.array([[1, 2, 3], [4, 5, 6]])
result = arr.sum(axis=0)
```

a) [5, 7, 9] (Sum of columns)

86. Given the following DataFrame:

```
import pandas as pd
df = pd.DataFrame({'A': [1, 2, 3], 'B': [4, 5, 6]})
df['C'] = df['A'] + df['B']
print(df)
```

a)

```
   A  B  C
0  1  4  5
1  2  5  7
2  3  6  9
```

87. Given the following DataFrame:

```
import pandas as pd
df = pd.DataFrame({'A': [1, 2, 3], 'B': [4, 5, 6]})
df['C'] = df['A'].astype(str) + df['B'].astype(str)
print(df)
```

a)

```
   A  B   C
0  1  4  14
1  2  5  25
2  3  6  36
```

88. Given the following DataFrame:

```
import pandas as pd
df = pd.DataFrame({'A': [1, 2, 3], 'B': [4, 5, 6]})
df = df[df['A'] > 1]
print(df)
```

c)

```
   A  B
1  2  5
2  3  6
```

89. Given the following DataFrame:

```
import pandas as pd
df = pd.DataFrame({'A': [1, 2, 3], 'B': [4, 5, 6]})
df['C'] = df['A'] * 2
df.set_index('C', inplace=True)
print(df)
```

a)

```
   A  B
C
2  1  4
4  2  5
6  3  6
```

90. Given the following DataFrame:

```
import pandas as pd
df = pd.DataFrame({'A': [1, 2, 3], 'B': [4, 5, 6]})
df['C'] = df['A'].apply(lambda x: x + 1)
df.drop(columns=['B'], inplace=True)
```

print(df)

a)

A C

0 1 2

1 2 3

2 3 4

91. Complete the code to create a simple line plot:

```
import matplotlib.pyplot as plt
x = [1, 2, 3, 4, 5]
y = [1, 4, 9, 16, 25]
plt.plot(x, y, 'b-') # Line color: blue, line style: solid
plt.xlabel('X-axis')
plt.ylabel('Y-axis')
plt.title('Simple Line Plot')
plt. ____
```

a) plt.show()

92. Given the following code, what will be the output?

```
import matplotlib.pyplot as plt
labels = ['A', 'B', 'C']
sizes = [15, 30, 55]
plt.pie(sizes, labels=labels, autopct='%1.1f%%', startangle=90)
plt.axis('equal')
plt.show()
```

b) A pie chart with three slices

93. Complete the code to create a histogram with 20 bins using Matplotlib:

```
import matplotlib.pyplot as plt
import numpy as np
data = np.random.randn(1000)
plt.hist(data, _____)
```

Ans: 20 (you would write 20 as the number of bins).

94. What will be the output of the following code?

```
import matplotlib.pyplot as plt
x = [1, 2, 3]
y = [4, 5, 6]
plt.plot(x, y, marker='o', linestyle='--', color='r')
plt.show()
```

a) A red dashed line plot with circular markers

95. Complete the following code to create a subplot with 1 row and 2 columns, displaying a scatter plot in the first subplot.

```
import matplotlib.pyplot as plt
x = [1, 2, 3, 4, 5]
y = [2, 3, 5, 7, 11]
fig, axs = plt.subplots(1, 2)
axs[0].scatter(x, y)
axs[1]._____
plt.show()
```

b) plot(x, y)

96. What will be the output of the following code? import matplotlib.pyplot as plt

```
labels = ['A', 'B', 'C']
sizes = [10, 20, 30]
plt.bar(labels, sizes)
plt.show()
```

a) A bar chart with three bars

97. Complete the code to create a figure with a size of 10x6 inches in Matplotlib.

```
import matplotlib.pyplot as plt
plt.figure(figsize=(_____, 6))
plt.plot([1, 2, 3], [4, 5, 6])
plt.show()
```

b) 10

98. What is the output of the following code?

```
import matplotlib.pyplot as plt
plt.plot([1, 2, 3], [4, 5, 6], 'ro')
plt.show()
```

b) A scatter plot with red circular markers

99. Complete the code to create a horizontal bar plot with Matplotlib.

```
import matplotlib.pyplot as plt
categories = ['A', 'B', 'C']
values = [5, 10, 15]
plt.barh(categories, values)
plt.show()
```

a) plt.barh()

100. What will be the output of the following code?

```
import matplotlib.pyplot as plt
plt.plot([1, 2, 3], [4, 5, 6], linestyle='--', color='g')
plt.show()
```

a) A green dashed line plot

101. Complete the code to set the y-axis limit from 0 to 10.

```
import matplotlib.pyplot as plt
plt.plot([1, 2, 3], [4, 5, 6])
plt.ylim(0, 10)
plt.show()
```

a) 0

102. What will be the output of the following code?

```
import matplotlib.pyplot as plt
plt.scatter([1, 2, 3], [4, 5, 6], c='b', marker='x')
plt.show()
```

a) A scatter plot with blue 'x' markers

103. Complete the code to add a grid to the scatter plot.

```
import matplotlib.pyplot as plt
plt.scatter([1, 2, 3], [4, 5, 6])
plt.grid()
plt.show()
```

b) plt.grid()

104. What will be the output of the following code?

```
import matplotlib.pyplot as plt
plt.plot([1, 2, 3], [4, 5, 6], linestyle='--', linewidth=2)
plt.show()
```

a) A dashed line plot with a width of 2

105. Complete the code to create a pie chart with percentages displayed inside the slices.

```
import matplotlib.pyplot as plt
sizes = [15, 30, 55]
labels = ['A', 'B', 'C']
plt.pie(sizes, labels=labels, autopct='_____%')
plt.show()
```

a) '%1.1f%%'