

Pandas Top 50 Cheat Sheet

Pandas Top 50 Cheat Sheet

1. How do you import pandas?

```
import pandas as pd
```

2. How do you create a DataFrame from a dictionary?

```
data = {'column1': [1, 2, 3], 'column2': ['a', 'b', 'c']}
```

```
df = pd.DataFrame(data)
```

3. How do you read a CSV file into a DataFrame?

```
df = pd.read_csv('file.csv')
```

4. How do you display the first few rows of a DataFrame?

```
df.head()
```

5. How do you display the last few rows of a DataFrame?

```
df.tail()
```

6. How do you get a random sample of rows from a DataFrame?

```
df.sample(n=5)
```

7. How do you get a summary of the DataFrame?

```
df.info()
```

Pandas Top 50 Cheat Sheet

8. How do you get descriptive statistics of the DataFrame?

```
df.describe()
```

9. How do you select a single column from a DataFrame?

```
df['column1']
```

10. How do you select multiple columns from a DataFrame?

```
df[['column1', 'column2']]
```

11. How do you select rows by their position?

```
df.iloc[0:5]
```

12. How do you select rows by their labels?

```
df.loc[0]
```

13. How do you filter rows based on a condition?

```
df[df['column1'] > 2]
```

14. How do you add a new column to a DataFrame?

```
df['new_column'] = df['column1'] + df['column2']
```

15. How do you drop a column from a DataFrame?

```
df.drop('column1', axis=1, inplace=True)
```

16. How do you rename columns in a DataFrame?

Pandas Top 50 Cheat Sheet

`df.rename(columns={'old_name': 'new_name'}, inplace=True)`

17. How do you check for missing data?

`df.isnull()`

18. How do you count missing values in each column?

`df.isnull().sum()`

19. How do you drop rows with missing values?

`df.dropna()`

20. How do you fill missing values with a specific value?

`df.fillna(value=0)`

21. How do you group data by a column and calculate the mean of each group?

`df.groupby('column1').mean()`

22. How do you perform aggregation on grouped data?

`df.groupby('column1').agg({'column2': 'sum', 'column3': 'mean'})`

23. How do you concatenate DataFrames vertically?

`pd.concat([df1, df2], axis=0)`

24. How do you concatenate DataFrames horizontally?

`pd.concat([df1, df2], axis=1)`

Pandas Top 50 Cheat Sheet

25. How do you merge two DataFrames on a key column?

```
pd.merge(df1, df2, on='key_column')
```

26. How do you save a DataFrame to a CSV file?

```
df.to_csv('file.csv', index=False)
```

27. How do you save a DataFrame to an Excel file?

```
df.to_excel('file.xlsx', index=False)
```

28. How do you sort a DataFrame by a column?

```
df.sort_values(by='column1')
```

29. How do you check for duplicate rows?

```
df.duplicated()
```

30. How do you drop duplicate rows?

```
df.drop_duplicates()
```

31. How do you apply a function to a column?

```
df['new_column'] = df['column1'].apply(lambda x: x * 2)
```

32. How do you reset the index of a DataFrame?

```
df.reset_index(drop=True, inplace=True)
```

Pandas Top 50 Cheat Sheet

33. How do you set a new index for a DataFrame?

```
df.set_index('column1', inplace=True)
```

34. How do you create a pivot table?

```
df.pivot_table(values='column1', index='column2', columns='column3', aggfunc='mean')
```

35. How do you melt a DataFrame from wide to long format?

```
pd.melt(df, id_vars=['column1'], value_vars=['column2', 'column3'])
```

36. How do you check the data types of each column?

```
df.dtypes
```

37. How do you change the data type of a column?

```
df['column1'] = df['column1'].astype('float')
```

38. How do you create a new DataFrame with only numeric columns?

```
df.select_dtypes(include=[np.number])
```

39. How do you create a copy of a DataFrame?

```
df_copy = df.copy()
```

40. How do you filter columns based on a condition?

```
df.loc[:, df.columns.str.contains('column')]
```

41. How do you find the unique values in a column?

Pandas Top 50 Cheat Sheet

`df['column1'].unique()`

42. How do you count the frequency of unique values in a column?

`df['column1'].value_counts()`

43. How do you create dummy variables for categorical data?

`pd.get_dummies(df, columns=['column1'])`

44. How do you concatenate strings in a column?

`df['new_column'] = df['column1'] + ' ' + df['column2']`

45. How do you find the correlation between columns?

`df.corr()`

46. How do you apply a function to every element in a DataFrame?

`df.applymap(lambda x: x * 2)`

47. How do you drop rows based on a condition?

`df.drop(df[df['column1'] < 2].index, inplace=True)`

48. How do you create a DataFrame from a list of lists?

`data = [[1, 'a'], [2, 'b'], [3, 'c']]`

`df = pd.DataFrame(data, columns=['column1', 'column2'])`

49. How do you create a DataFrame from a NumPy array?

Pandas Top 50 Cheat Sheet

```
import numpy as np
```

```
data = np.array([[1, 2], [3, 4]])
```

```
df = pd.DataFrame(data, columns=['column1', 'column2'])
```