Pandas Cheat Sheet for Data Science

|  |  |
| --- | --- |
| Pandas Topic | Code Example |
| Install Pandas | pip install pandas |
| Import Pandas | import pandas as pd |
| Method Chaining | df = (pd.melt(df).rename(columns={'variable': 'var'}).query('val >= 200')) |
| Pandas Syntax | df[(df['col\_1']=='a')&(df['col\_2']>=10)] |
| Create new Series | s = pd.Series(['a', 'b', 'c'], index=[0, 1, 2]) |
| Create new DataFrame | df = pd.DataFrame({'col\_1': [11, 12, 13], 'col\_2': [21, 22, 23], 'col\_3': [31, 32, 33]}, index=[0, 1, 2]) |
| Read from CSV | pd.read\_csv(filename) |
| Read from Excel | pd.read\_excel(filename) |
| Read from SQL | pd.read\_sql(query, connection\_object) |
| Read from JSON | pd.read\_json(json\_string) |
| Write to CSV | df.to\_csv(filename) |
| Write to Excel | df.to\_excel(filename) |
| Write to JSON | df.to\_json(filename) |
| Write to HTML | df.to\_html(filename) |
| View first n rows | df.head(n) |
| View last n rows | df.tail(n) |
| Get shape | df.shape |
| Info summary | df.info() |
| Describe summary | df.describe() |
| View unique counts | s.value\_counts(dropna=False) |
| Select by index | s.loc[0] |
| Select by position | s.iloc[0] |
| Select single column | df['col\_1'] |
| Select multiple columns | df[['col\_1', 'col\_2']] |
| Select rows meeting condition | df.loc[df['col\_1'] > 10, ['col\_1', 'col\_2']] |
| Add new column | df['new col'] = df['col'] \* 100 |
| Add new row | df.loc[-1] = [1, 2, 3] |
| Drop column | df.drop('col\_1', axis=1) |
| Drop NaN rows | df.dropna() |
| Sort values | df.sort\_values(by='col\_1') |
| Reset index | df.reset\_index() |
| Filter rows | df[df['col\_1'] > 100] |
| Group by single column | df.groupby('col\_1') |
| Convert to string | df['points'].astype(str) |
| Convert to int | df['col\_1'].astype('int64') |
| Convert to float | df['col\_1'].astype(float) |
| Merge DataFrames | df1.join(df2, on=col1, how='inner') |
| Concat DataFrames | pd.concat([df1, df2], axis=1) |
| Apply function | df.apply(calc, axis=1) |
| Lambda apply | df.apply(lambda x: x \* -1 if x < 0 else x) |