Pandas and NumPy Interview Questions

Pandas Interview Questions

Q: What is Pandas?

A: Pandas is a Python library for data manipulation and analysis. It provides Series and DataFrame structures.

Q: Difference between Series and DataFrame?

A: Series is 1D; DataFrame is 2D with columns of potentially different types.

Q: How to read/write CSV?

A: pd.read_csv('file.csv'); df.to_csv('out.csv')

Q: How to get info about a DataFrame?

A: df.info(), df.describe(), df.shape

Q: Selecting columns/rows?

A: df['col'], df.loc[0], df.iloc[0]

Q: Filtering rows?

A: df[df['age'] > 30]

Q: Handling missing data?

A: df.isnull().sum(), df.dropna(), df.fillna(0)
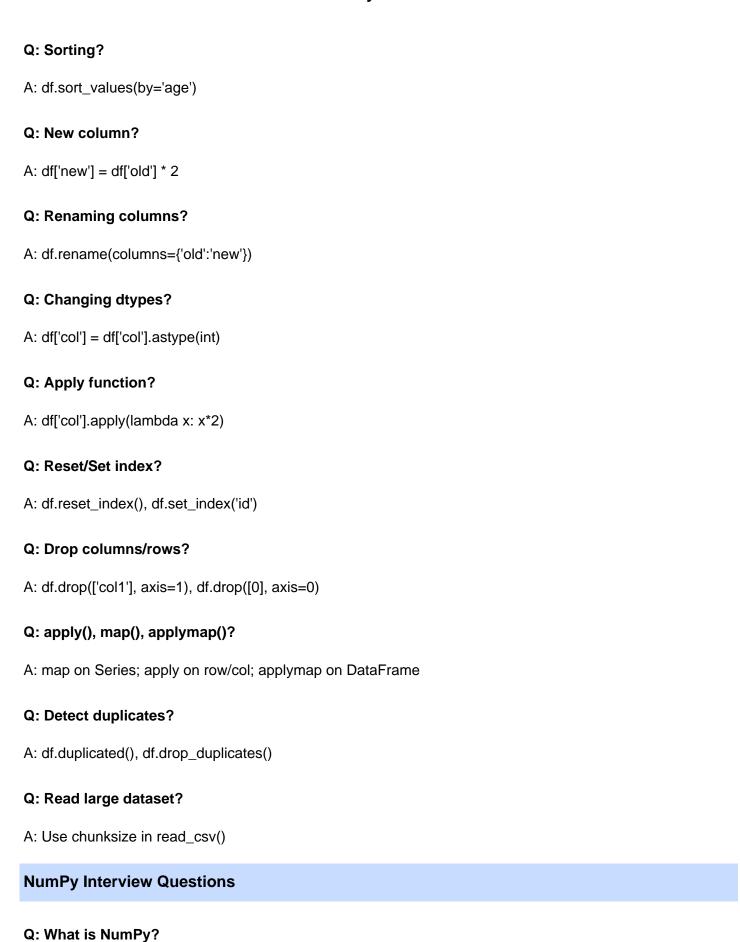
Q: Using groupby?

A: df.groupby('dept')['salary'].mean()

Q: Merging DataFrames?

A: pd.merge(df1, df2, on='id')

Q: Pivot table?

A: df.pivot_table(index='dept', columns='gender', values='salary')

Pandas and NumPy Interview Questions

Q: Sorting?

A: df.sort_values(by='age')

Q: New column?

A: df['new'] = df['old'] * 2

Q: Renaming columns?

A: df.rename(columns={'old':'new'})

Q: Changing dtypes?

A: df['col'] = df['col'].astype(int)

Q: Apply function?

A: df['col'].apply(lambda x: x*2)

Q: Reset/Set index?

A: df.reset_index(), df.set_index('id')

Q: Drop columns/rows?

A: df.drop(['col1'], axis=1), df.drop([0], axis=0)

Q: apply(), map(), applymap()?

A: map on Series; apply on row/col; applymap on DataFrame

Q: Detect duplicates?

A: df.duplicated(), df.drop_duplicates()

Q: Read large dataset?

A: Use chunksize in read_csv()

NumPy Interview Questions

Q: What is NumPy?

A: A library for fast numerical computation and array operations.

Q: What is ndarray?

A: A multi-dimensional array of same-type elements.

Q: Creating arrays?

A: np.array, np.zeros, np.ones, np.arange, np.linspace

Q: Array attributes?

A: shape, ndim, dtype, size

Q: Reshape array?

A: a.reshape((2,3))

Q: reshape() vs resize()?

A: reshape returns a view; resize modifies original

Q: Basic operations?

A: a+b, a*b, np.sum(a), np.mean(a)

Q: Indexing/slicing?

A: a[0], a[:3], a[::2]

Q: Broadcasting?

A: Enables arithmetic on arrays of different shapes

Q: Matrix multiplication?

A: np.dot(a, b), a @ b

Q: Random numbers?

A: np.random.rand, randint, randn

Q: Flatten array?

A: a.flatten(), a.ravel()

Q: Concatenate arrays?

A: np.concatenate, vstack, hstack

Q: Handle NaNs?

A: np.isnan, np.nan_to_num, np.nanmean

Q: copy() vs assignment?

A: copy() makes deep copy

Q: Unique elements?

A: np.unique(arr, return_counts=True)

Q: Element-wise max/min?

A: np.maximum(a, b), np.minimum(a, b)

Q: Boolean masking?

A: a[a > 5]

Q: Sort arrays?

A: np.sort(a), np.argsort(a)

Q: Save/load arrays?

A: np.save, np.load