



--	--	--	--	--	--	--	--	--	--

Question Paper Version : A

Fourth Semester B.E. Degree Examination, Dec.2023/Jan.2024
Data Cleaning and Preparation with Python Pandas

[Max. Marks: 50

INSTRUCTIONS TO THE CANDIDATES

1. Answer all the **fifty** questions, each question carries one mark.
2. Use only **Black ball point pen** for writing / darkening the circles.
3. **For each question, after selecting your answer, darken the appropriate circle corresponding to the same question number on the OMR sheet.**
4. Darkening two circles for the same question makes the answer invalid.
5. **Damaging/overwriting, using whiteners** on the **OMR** sheets are strictly prohibited.

-
1. What is the primary data structure used for handling panel data in Pandas?
a) Series b) Data frame c) Panel d) Array
 2. Which function is used to create a Panel from 3D ndarray?
a) pd.Panel () b) pd.DataFrame ()
c) pd.Panel.from_dict d) pd.Panel.from_array ()
 3. How are items, major_axis and minor_axis represented in a Pandas Panel?
a) columns, Rows, values b) Rows, Columns, Values
c) Values, Rows, Columns d) Values, Columns, Rows
 4. In Pandas, which method is used to convert a panel to a DataFrame?
a) Panel.to_df () b) Panel.to_dataframe ()
c) pd.DataFrame.from_panel () d) Panel.to-frame ()
 5. Which attribute in a Pandas Panel returns the number of items along the axis 0?
a) panel.items b) panel.major_axis c) panel.minor_axis d) panel.shape
 6. Which method is used to perform mathematical operations between two panels in Pandas?
a) panel.add () b) panel.combine ()
c) panel.merge () d) panel.combine_first ()
 7. Which attribute in a Pandas panel returns the datatype of the values stored in the Panel?
a) panel.dtypes b) panel.data_type
c) panel.values.dtype d) panel.dtype

8. In Pandas, what method is used to fill missing values in a Panel object with a specified values?
a) panel.fill_na ()
b) panel.fill_missing ()
c) panel.fill ()
d) panel.fill_values ()
9. Which function is used to create a Panel by loading data from an Excel file in Pandas?
a) pd.Panel.read_excel ()
b) pd.read_panel_excel ()
c) pd.read_excel_panel ()
d) pd.read_excel ()
10. In Pandas, what method is used to reset the index of a Data Frame to the default integer index?
a) reset_index ()
b) reindex ()
c) set_default_index
d) index_reset ()
11. What is used to create a MultiIndex in Pandas?
a) pd.MultiIndex ()
b) create_MultiIndex ()
c) pd.Index
d) MultiIndex.Create ()
12. Which method is used to access data from a specific level of a MultiIndex in Pandas?
a) get_level_values ()
b) level_values ()
c) select_level ()
d) access_level ()
13. In Pandas MultiIndex, which method is used to perform a partial indexing at inner levels of the index?
a) slice ()
b) partial_index ()
c) partial_select ()
d) xs ()
14. Which method in Pandas is used to sort levels of a MultiIndex along a particular level?
a) sort_index ()
b) sort_levels ()
c) sort ()
d) sort_multiindex ()
15. In Pandas, what method is used to create a DataFrame with a hierarchical index from a dictionary of arrays?
a) pd.DataFrame.from_dict ()
b) pd.MultiIndex.from_dict ()
c) pd.DataFrame.from_multiindex ()
d) pd.MultiIndex.from_arrays ()
16. What method in Pandas is used to create a MultiIndex from product of input iterables?
a) pd.MultiIndex.from_tuples ()
b) pd.MultiIndex.from_product ()
c) pd.MultiIndex.create_index ()
d) pd.MultiIndex.create_product ()
17. Which Pandas function is used to concatenate two or more DataFrames along rows or columns?
a) pd.concat ()
b) pd.merge ()
c) pd.join ()
d) pd.append ()
18. Which Pandas function is used to reshape a DataFrame using pivot tables?
a) pd.pivot_table ()
b) pd.melt ()
c) pd.pivot ()
d) pd.reshape_table ()
19. What function is used to compute pairwise counts between two or more categorical variables in Pandas?
a) pd.count_table ()
b) pd.crosstab ()
c) pd.pivot_table ()
d) pd.pairwise_counts ()

20. Which method is used in Pandas to stack one level of a DataFrame's column labels to produce a MultiIndex?
 a) df.stack () b) df.unstack () c) df.pivot () d) df.melt ()
21. Which Pandas method is used to check if a string contains a specified substring within a series?
 a) contains () b) find ()
 c) match () d) search ()
22. Which method in Pandas series is used to convert strings to lowercase?
 a) to_lowercase () b) lowercase ()
 c) str.lower () d) convert_to_lower ()
23. Which method in Pandas is used to split each string in a series by a specified delimiter?
 a) str.split () b) split_string ()
 c) split () d) str.extractall ()
24. In Pandas, which function is used to convert a series of strings to categorical values?
 a) pd.to_categorical () b) pd.convert_to_categorical ()
 c) pd.categorical () d) pd.to_numeric ()
25. In Pandas, which method is used to find the index of the first occurrence of a substring in each element of a series?
 a) str.find () b) str.index () c) str.locate () d) str.first_occurrence ()
26. Which Pandas method is used to detect missing values in a DataFrame?
 a) detect_missing () b) is_null ()
 c) has_missing () d) isnull ()
27. In Pandas, what method is used to interpolate missing values in a DataFrame using various interpolation methods?
 a) interpolate () b) fill_missing ()
 c) interpolate_values d) impute ()
28. Which Pandas function is used to create a Boolean mask of non-missing values in a DataFrame?
 a) isna () b) notnull () c) non_missing () d) valid_values ()
29. What method in Pandas is used to check if a DataFrame has any missing values?
 a) missing () b) check_missing () c) isnull () d) any_missing ()
30. What Pandas function is used to replace missing values in a DataFrame with the mean of the column?
 a) replace_mean () b) fill_mean ()
 c) mean_replace () d) fillna () with df.mean ()
31. Which Pandas method is used to create groups based on a specified criterion?
 a) group () b) create_group ()
 c) groupby () d) grouping ()

32. What is the purpose of the `size ()` method in Pandas groupby object?
a) counts the number of rows in each group.
b) splits the DataFrame into multiple groups
c) calculates the mean of each group
d) renames the group indices
33. In Pandas groupby object, which method is used to iterate through the groups?
a) loop b) `for_groups ()` c) `group_iter ()` d) `_iter_ ()`
34. What Pandas method is used to retrieve a specific group from a groupby object by key?
a) `get_group ()` b) `retrieve_group ()`
c) `select_group ()` d) `group_select ()`
35. In Pandas groupby object, which method is used to apply a function to each group individually?
a) `apply ()` b) `for_each ()`
c) `process ()` d) `group_apply ()`
36. Which method in Pandas is used to return the unique group keys from a groupby object?
a) `unique_keys ()` b) `group_keys ()`
c) `keys ()` d) `groups ()`
37. What is the purpose of the `transform ()` method in a Pandas groupby object?
a) Filters the groups based on a condition
b) Applies a function to each group individually
c) Performs aggregation on the groups
d) Broadcasts the aggregated results back to the original DataFrame shape.
38. What Pandas method is used to remove one or more groups from a groupby object based on group keys?
a) `remove_group ()` b) `group_remove ()`
c) `drop_group ()` d) `filter ()`
39. In Pandas Groupby object, what method is used to rank values within groups?
a) `rank ()` b) `group_rank ()`
c) `groupby_rank ()` d) `grouped_rank ()`
40. What Pandas method is used to pivot the groups into columns?
a) `pivot ()` b) `unstuck ()` c) `spread ()` d) `group_pivot ()`
41. Which Pandas method is used to convert a column to datetime format in a DataFrame?
a) `to_time ()` b) `convert_datetime ()`
c) `pd.datetime ()` d) `pd.to_datetime ()`
42. What Pandas method is used to create a time delta object representing a duration?
a) `timedelta ()` b) `create_timedelta ()`
c) `pd.timedelta ()` d) `pd.Timedelta ()`
43. In Pandas, which function is used to plot data directly from a DataFrame?
a) `plot ()` b) `visualize ()` c) `pd.plot ()` d) `df.plot ()`

- 44. Which Pandas method is used to efficiently handle large datasets by loading data in chunks?
a) `read_chunk ()` b) `read_large ()`
c) `chunk ()` d) `read_csv ()` with chunksize

- 45. What Pandas function is used to resample time series data to a different frequency?
a) `resample ()` b) `change_frequency ()`
c) `time_resample ()` d) `pd.resample ()`

- 46. Which method in Pandas is used to shift the index by a specified number of periods?
a) `shift ()` b) `move_index ()`
c) `index_shift ()` d) `pd.shift_index ()`

- 47. What method in Pandas is used to interpolate missing values in time series data?
a) `time_interpolate ()` b) `interpolate ()`
c) `fill_missing ()` d) `time_fillna ()`

- 48. In Pandas, what function is used to handle time zone conversions for time series data?
a) `convert_timezone ()` b) `tz_convert ()`
c) `timezone_convert ()` d) `pd.tz_convert ()`

- 49. What method in Pandas DataFrame is used to shift the values of a DataFrame by a specified number of periods?
a) `shift ()` b) `move_values ()` c) `transpose ()` d) `pd.shift_values ()`

- 50. Which method in Pandas DataFrame is used to generate descriptive statistics of time series data?
a) `describe ()` b) `stats ()`
c) `summary_stats ()` d) `time_describe ()`
