

## 21CSC554J Applied Data Science Lab

### Lab 3 – Indexing and Reindexing in pandas dataframes

#### Question 1 — label vs position selection

loc (label) vs iloc (position).

- Load nba.csv with Name as index.
- Print the row for player "Avery Bradley" using .loc, then print the 4th row in the DataFrame using .iloc.
- Using .loc, extract the "Team" and "Position" for "R.J. Hunter".

What to use: `pd.read_csv(..., index_col="Name")`, `.loc`, `.iloc`.

Hints:

`.loc["Avery Bradley"]` returns a Series of that player's data

`.iloc[3]` returns the 4th row by position

`data.loc["R.J. Hunter", ["Team", "Position"]]` returns a two-element Series or 1x2 DataFrame.

#### Question 2 — set\_index / reset\_index and at/iat

Setting and resetting index and accessing single values.

- Read nba.csv (no index\_col this time).
- Set the DataFrame index to the "Team" column (use `set_index`) and then reset it back to default index.
- Using the original DataFrame (with Name as index), retrieve the Salary of "Avery Bradley" using both `.at` and `.iat` (use `.at` with label and `.iat` with integer position).

What to use: `df.set_index("Team", inplace=...)`, `df.reset_index()`, `.at`, `.iat`.

Hints:

After `set_index("Team")`, the index values are team names. `.at["Avery Bradley", "Salary"]` gives label-based scalar

to use `.iat` you must know the integer row/col positions (e.g., find row number with `.get_loc` on the index or use `.reset_index()` first).

#### Question 3 — reindex rows to add players and fill values

Application of Reindexing: Reindex to add missing labels and fill missing data.

- Load nba.csv with `index_col="Name"`.
- Create a new index list that includes the existing first 10 player names plus two new names ["New Player A", "New Player B"].
- Reindex the DataFrame to that list and fill missing numeric fields with 0 and missing string fields with "Unknown".
- Show the resulting rows for the two new players.

What to use: `df.reindex(new_index, fill_value)` or `df.reindex(...).fillna(...)`.

Hints:

The two added player rows should appear with NaN replaced (e.g., Salary = 0, College = "Unknown"). Explain which approach you used (single `fill_value` works only for entire DataFrame; `fillna` with a dict is more flexible).

#### **Question 4 — reorder / add columns with reindex (columns axis)**

Reindex columns to reorder and introduce a new column.

- Using `nba.csv` (index by Name), reindex the columns to: ["Team", "Position", "Salary", "College", "Height"].
- Ensure the reindexed DataFrame shows the requested column order and that any missing column is created with -1.

What to use: `df.reindex(columns=[...], fill_value=-1)` or `df.reindex(columns=...).assign(...)`.

Hints The output should have columns in the requested order; missing columns must be present and filled with -1 (numeric) so downstream numeric ops won't fail.