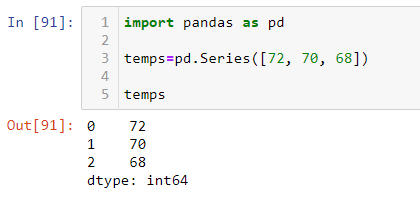
# Pandas

# Data Structures

1. Series: 1-demensional array of tabular data



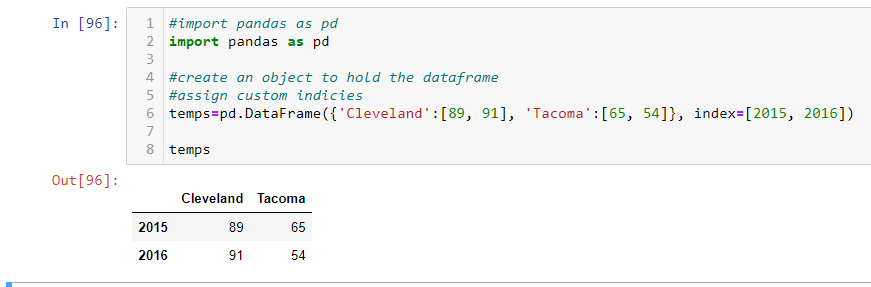
1. Data Frames: 2-demensional array. a collection of Series
   1. Can be created in the dataframe statement or an object can be passed as an argument
   2. It can take many forms, including a dictionary, list of dictionaries, and a Numpy ndarray,

# 

See further introduction at: <https://pandas.pydata.org/docs/user_guide/dsintro.html>

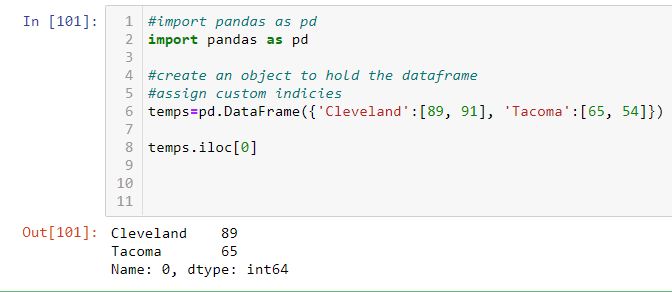
# Indexing

* Indices can be 0-based integers, or customized strings, integers, and floats
* If no index is assigned, 0-based integers are default

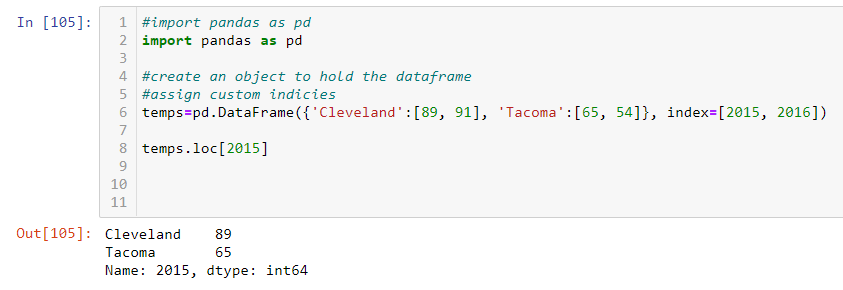


# Indexing rows and columns, using iloc and loc

* Index rows with default 0-based integer labels using iloc



* Index rows with custom integer labels that exceed the length of the axis using loc

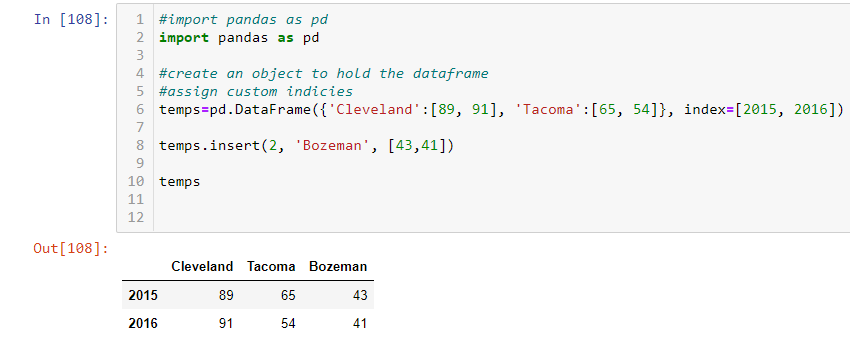


* Index rows with string labels using loc
* Index integer or string column labels using the structure: dataframe[int or ‘string’]

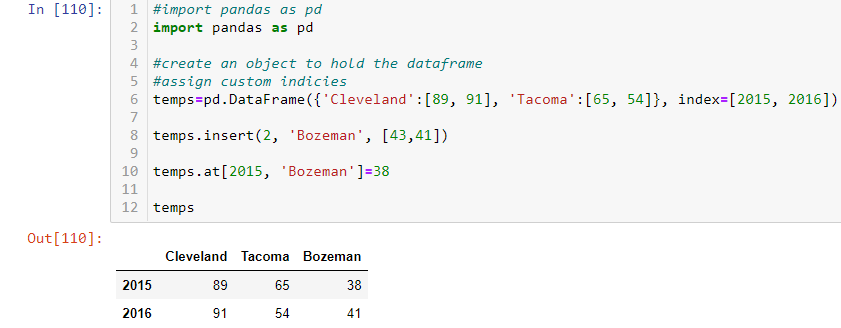
See more on indexing: <https://pandas.pydata.org/docs/reference/api/pandas.DataFrame.iloc.html>

Manipulation

* Add a column to a data frame in a specific location using the structure: dataframe.insert(position, column name, [value 1, value 2, etc.])



* Change the value of a single cell using the structure: dataframe.at[row, column]=value



# I/O

* Read a json file into a dataframe using: object=pd.read\_json(file\_name.json)

See more on I/O processing in Pandas: <https://pandas.pydata.org/pandas-docs/stable/user_guide/io.html#excel-files>