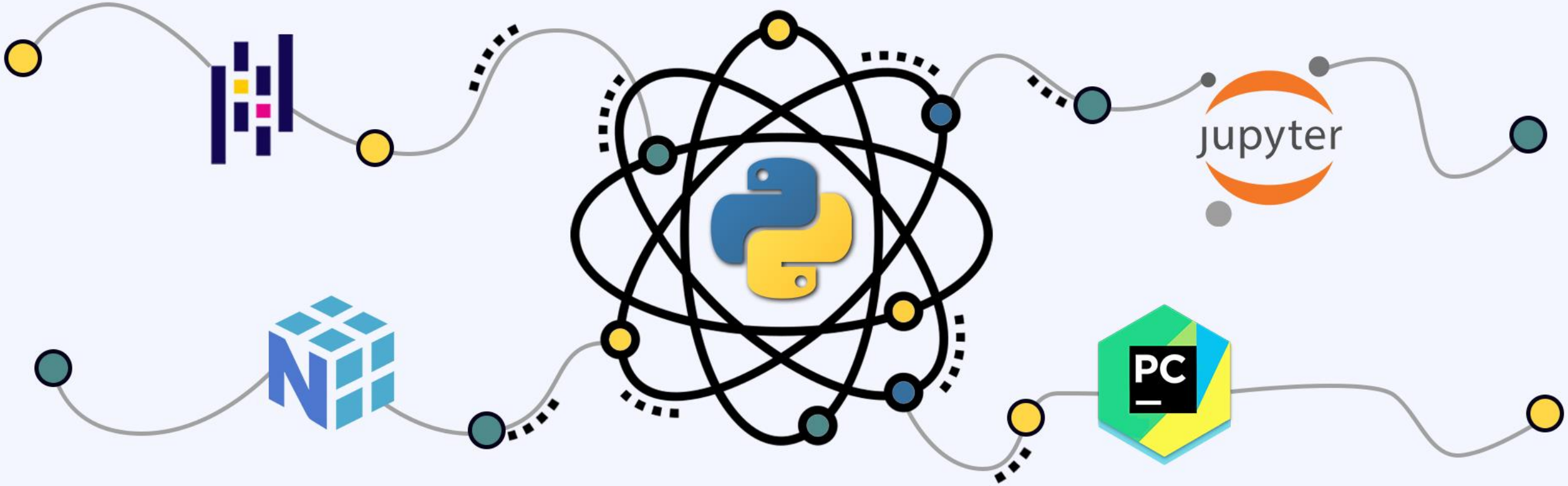
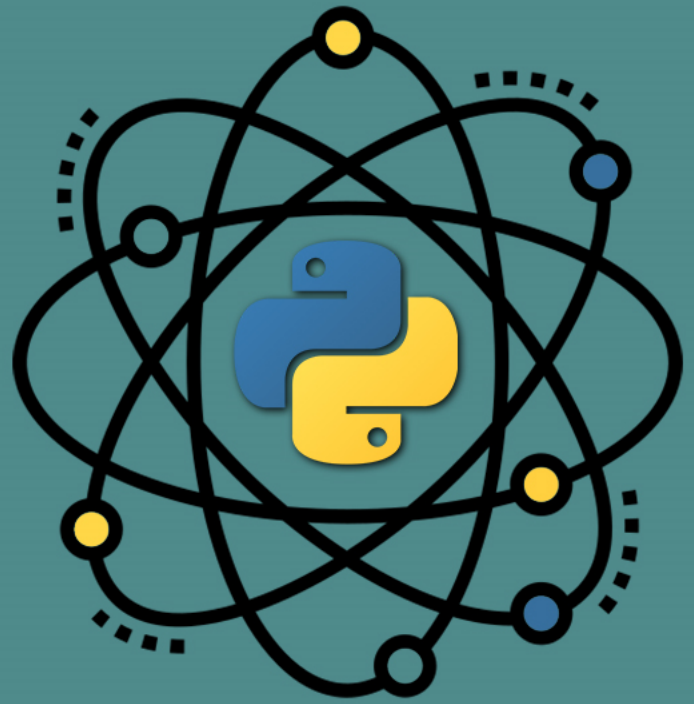


MOHAMMAD DEHGHANI PTYHON FOR DATA SCIENCE

زبان پایتون در علم داده

مدرس: محمد دهقانی



زبان پایتون در علم داده

مدرس:
محمد دهقانی

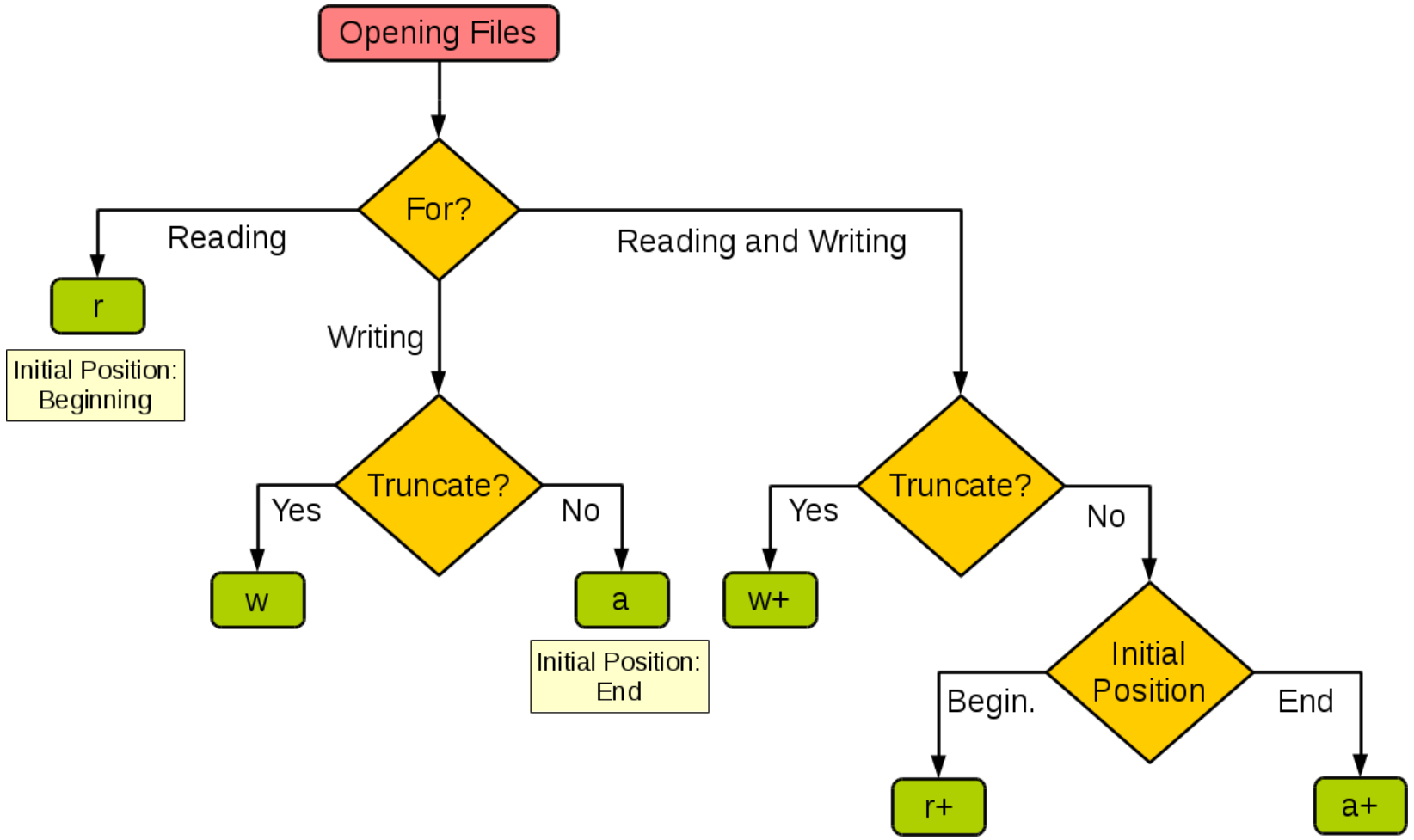
جلسه: هشتم
موضوع: فایل ها

مواردی که در این جلسه بررسی می شوند:

1. باز کردن فایل و نوشتن در آن
2. استفاده از with برای فایل ها
3. عملیات append در فایل ها
4. کار با pickle

Python File

An important component of an operating system is its file and directories.


Opening Files
For?
Reading
Writing
Reading and Writing
r
Initial Position: Beginning
Truncate?
Yes
No
w
a
Initial Position: End
Truncate?
Yes
No
w+
Initial Position
Begin.
End
r+
a+

Increase DataFrame by 1000x

from 1.2 MB to 1.2 GB

```
[ ]  1  bigger_file = pd.concat([price_data]*1000)
```

Save downloaded data to a CSV file

```
[ ]  1  %%timeit
     2  bigger_file.to_csv("price_data.csv")
```

```
1 loop, best of 3: 3min 58s per loop
```

Load data from CSV file

```
[ ]  1  %%timeit
     2  price_data_from_csv = pd.read_csv("price_data.csv", index_col=0, parse_dates=True)
```

```
1 loop, best of 3: 13.6 s per loop
```

Save data as pickle file

```
[52]  1  import pickle
      2
      3  def pickle_save(data, name):
      4      """Saves dataframe as pickle file into the current folder,
      5      takes name of df to be saved as first argument and filename
      6      (without '.pkl') as second argument
      7      """
      8      assert type(name) is str
      9      f = open(f"{name}.pkl", "wb")
     10      pickle.dump(data, f)
     11      f.close()
```

```
[ ]  1  %%timeit
     2  pickle_save(bigger_file, "data_pickle")
```

```
1 loop, best of 3: 15.8 s per loop
```

Load data from a pickle file

```
[54]  1  %%timeit
      2  with open("data_pickle.pkl", "rb") as f:
      3      price_data_from_pkl = pickle.load(f)
```

```
1 loop, best of 3: 419 ms per loop
```

source