Python Programming for Business Analytics

What is Business Analytics?

“Where do we come from? What are we? Where are we going?”

–Paul Gauguin

Where do we come from? What are we? Where are we going?

Artist: Paul Gauguin
Location: Museum of fine arts, Boston

Business Analytics

“Where do we come from?”

Business Analytics

“Where do we come from?”

Descriptive analytics:
Insights from historical data

- Business reporting
- Data visualization
- Data warehouse

Business Analytics

“Where do we come from?”

Descriptive analytics:
Insights from historical data

- Business reporting
- Data visualization
- Data warehouse

“What are we?”

Business Analytics

“Where do we come from?”

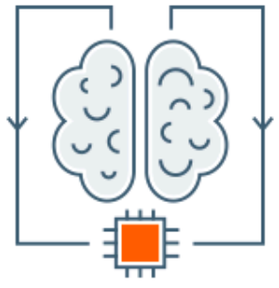
Descriptive analytics:
Insights from historical data

- Business reporting
- Data visualization
- Data warehouse

“What are we?”

Predictive analytics:
Current and future trends

- Statistics
- Econometrics
- Data mining
- Forecast
- Machine learning



Business Analytics

“Where do we come from?”

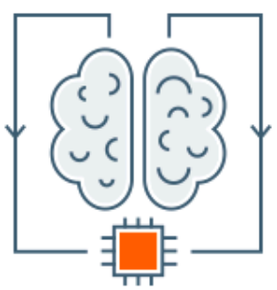
Descriptive analytics:
Insights from historical data

- Business reporting
- Data visualization
- Data warehouse

“What are we?”

Predictive analytics:
Current and future trends

- Statistics
- Econometrics
- Data mining
- Forecast
- Machine learning



“Where are we going?”

Business Analytics

“Where do we come from?”

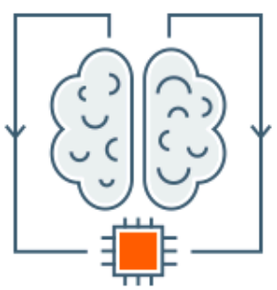
Descriptive analytics:
Insights from historical data

- Business reporting
- Data visualization
- Data warehouse

“What are we?”

Predictive analytics:
Current and future trends

- Statistics
- Econometrics
- Data mining
- Forecast
- Machine learning



“Where are we going?”

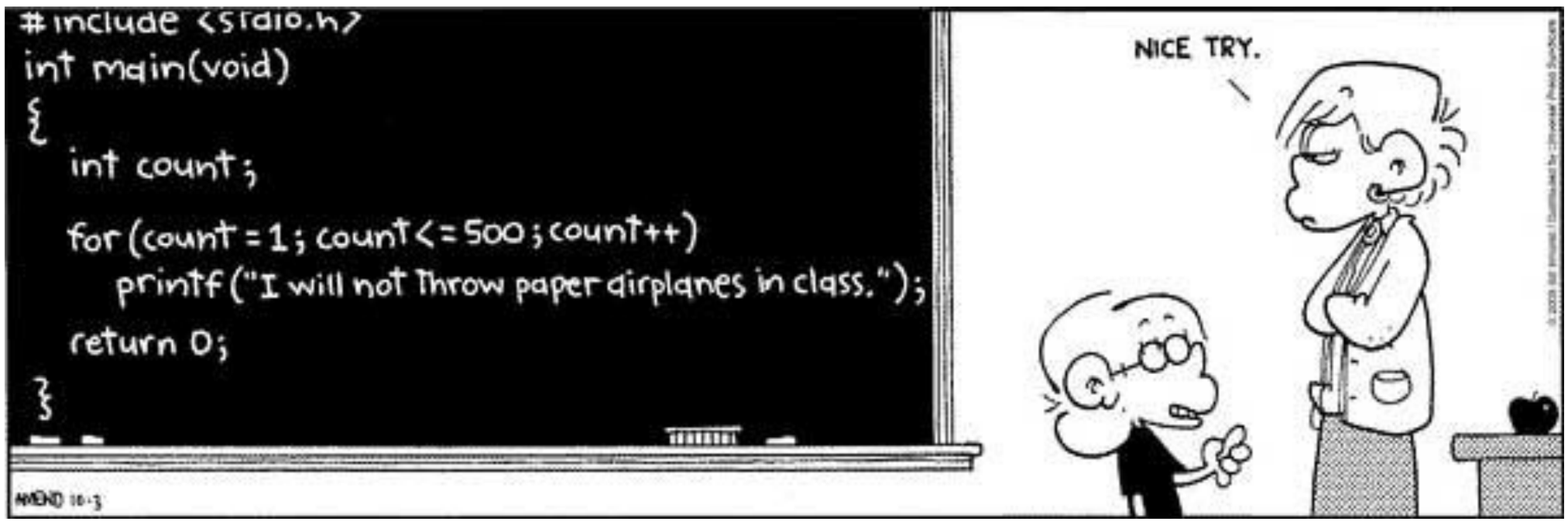
Prescriptive analytics:
Data-driven decision-making

- Optimization
- Decision tree
- Simulations
- Expert system

Why Python?

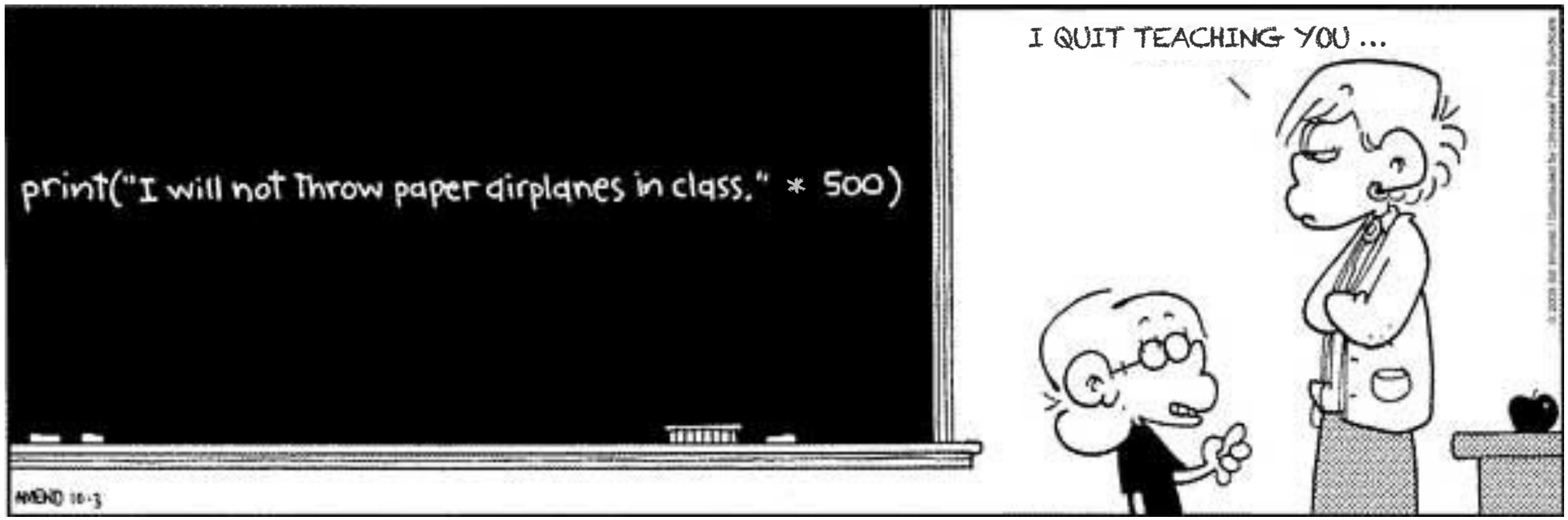
Python is the Best

- Evolution of programming
  - C language (1980s)



Python is the Best

- Evolution of programming
  - Python



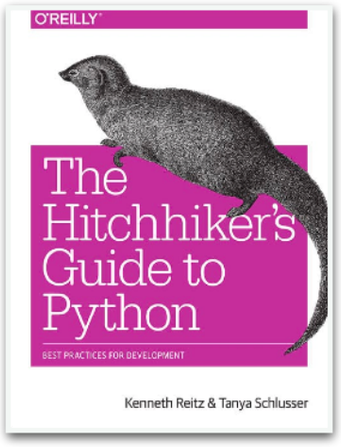
Python is the Best

- Evolution of programming
  - Python

> If you ask Python programmers what they like most about Python, they will often cite its high readability. Indeed, a high level of readability is at the heart of the design of the Python language, following the recognized fact that code is read much more often than it is written.

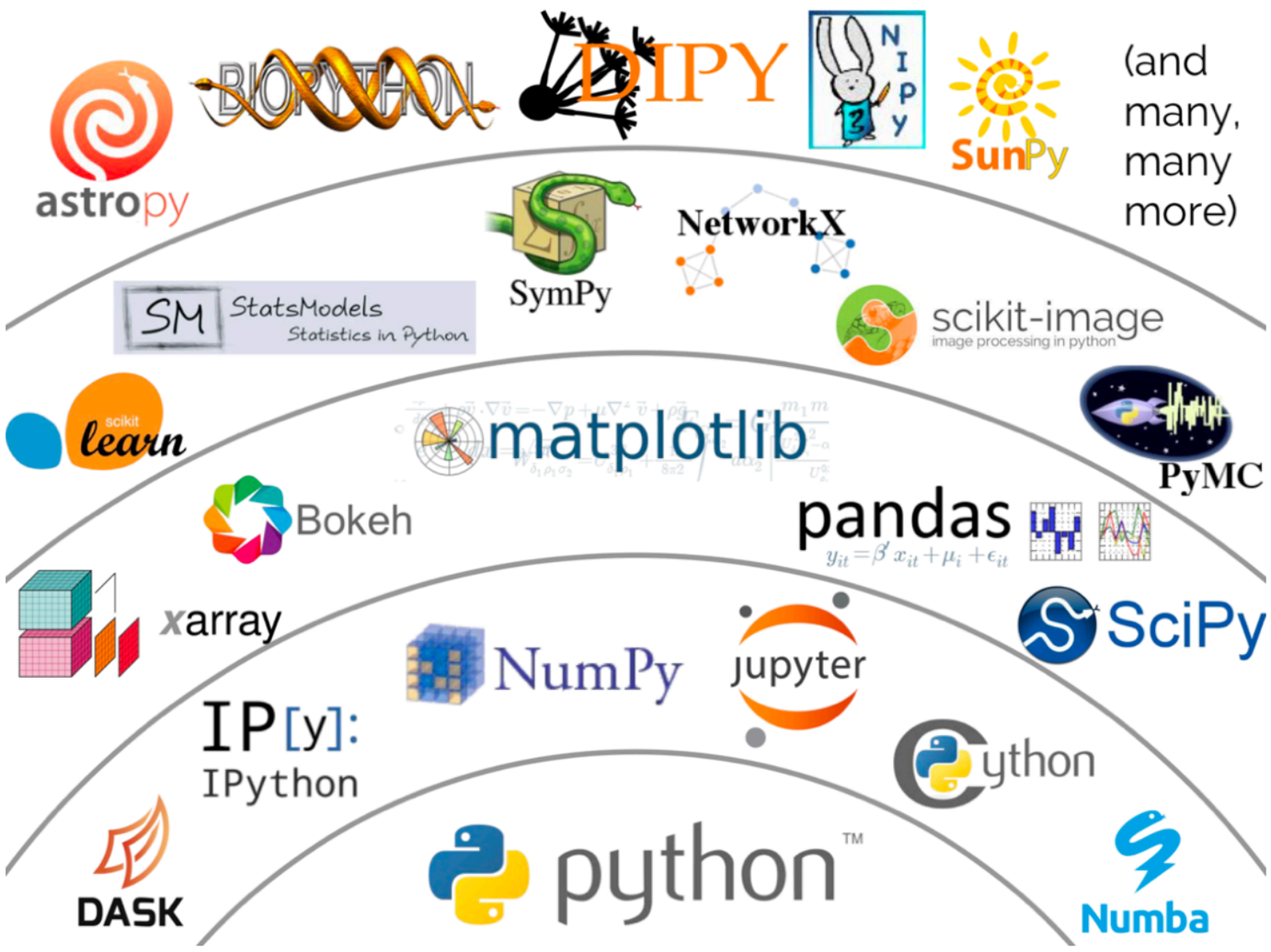
```
In [1]: print("I will not throw paper airplanes in class.\n" * 500)
```

```
I will not throw paper airplanes in class.
I will not throw paper airplanes in class.
I will not throw paper airplanes in class.
I will not throw paper airplanes in class.
```



Python is the Best

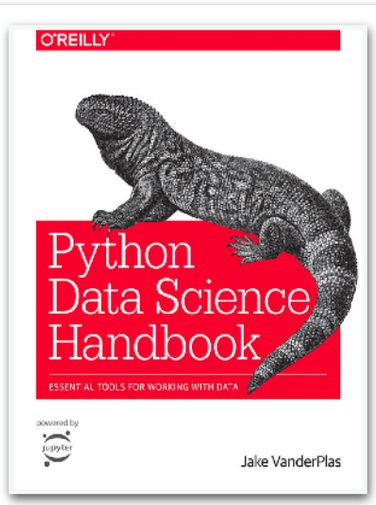
- The Python universe



Python is the Best

- The Python universe

The usefulness of Python for data science stems primarily from the large and active ecosystem of third-party packages: NumPy for manipulation of homogeneous array-based data, Pandas for manipulation of heterogeneous and labeled data, SciPy for common scientific computing tasks, Matplotlib for publication-quality visualizations, IPython for interactive execution and sharing of code, Scikit-Learn for machine learning, and many more tools that will be mentioned in the following pages.



Learning Materials

- Jupyter notebook files for learning and practice
- Jupyter notebook files for exercises
- Slides as supplementary materials

Data Visualization Applications

- Map of the blue death
- How the circle line rogue train was caught with data