

Data Science from Scratch

First Principles with Python

Stephen CUI 

March 20, 2023

Contents

Chapter 1

Getting Data

1.1 stdin and stdout

If you run your Python scripts at the command line, you can pipe data through them using `sys.stdin` and `sys.stdout`.

在Windows中的命令行可以输入：

```
type line_count.txt | python egrep.py "[0-9]"
type line_count.txt | python line_count.py
type line_count.txt | python egrep.py "[0-9]" | python line_count.py
```

The `|` is the pipe character, which means “use the output of the left command as the input of the right command.” You can build pretty elaborate data-processing pipelines this way.