

a

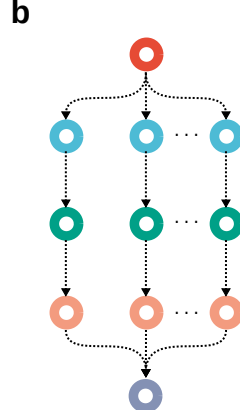
```

1  ● configfile: "config.yaml"
2
3  ● rule all:
4  ●   input:
5  ●       expand(
6  ●         "results/plots/{country}.hist.pdf",
7  ●         country=config["countries"]
8  ●       )
9
10 ● rule download_data:
11 ●   output:
12 ●       "data/worldcitiespop.csv"
13 ●   log:
14 ●       "logs/download.log"
15 ●   conda:
16 ●       "envs/curl.yaml"
17 ●   shell:
18 ●       "curl -L https://burntsushi.net/stuff/worldcitiespop.csv > {output} 2> {log}"
19
20 ● rule select_by_country:
21 ●   input:
22 ●       "data/worldcitiespop.csv"
23 ●   output:
24 ●       "results/by-country/{country}.csv"
25 ●   log:
26 ●       "logs/select-by-country/{country}.log"
27 ●   conda:
28 ●       "envs/xsv.yaml"
29 ●   shell:
30 ●       "xsv search -s Country '{wildcards.country}' "
31 ●       " {input} > {output} 2> {log}"
32
33 ● rule plot_histogram:
34 ●   input:
35 ●       "results/by-country/{country}.csv"
36 ●   output:
37 ●       "results/plots/{country}.hist.svg"
38 ●   container:
39 ●       "docker://faizanbashir/python-datascience:3.6"
40 ●   log:
41 ●       "logs/plot-hist/{country}.log"
42 ●   script:
43 ●       "scripts/plot-hist.py"
44
45 ● rule convert_to_pdf:
46 ●   input:
47 ●       "{prefix}.svg"
48 ●   output:
49 ●       "{prefix}.pdf"
50 ●   log:
51 ●       "logs/convert-to-pdf/{prefix}.log"
52 ●   wrapper:
53 ●       "0.47.0/utls/cairosvg"

```

Legend

- domain knowledge
- technical knowledge
- Snakemake knowledge
- trivial



c

```

import sys
sys.stderr = open(snakemake.log[0], "w")

import matplotlib.pyplot as plt
import pandas as pd

cities = pd.read_csv(snakemake.input[0])

plt.hist(cities["Population"], bins=50)

plt.savefig(snakemake.output[0])

```

