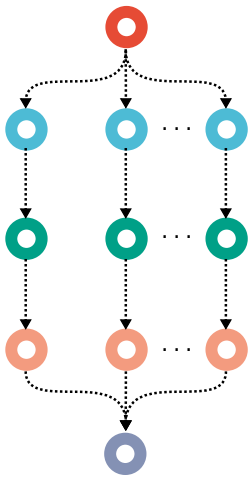
a

```
configfile: "config.yaml"

rule all:
    input:
        expand(
            "results/plots/{country}.hist.pdf",
            country=config["countries"]
        )

rule download_data:
    output:
        "data/worldcitiespop.csv"
    log:
        "logs/download.log"
    conda:
        "envs/curl.yaml"
    shell:
        "curl -L https://burntsushi.net/stuff/worldcitiespop.csv > {output} 2> {log}"

rule select_by_country:
    input:
        "data/worldcitiespop.csv"
    output:
        "results/by-country/{country}.csv"
    log:
        "logs/select-by-country/{country}.log"
    conda:
        "envs/xsv.yaml"
    shell:
        "xsv search -s Country '{wildcards.country}' "
        " {input} > {output} 2> {log}"

rule plot_histogram:
    input:
        "by-country/{country}.csv"
    output:
        "results/plots/{country}.hist.svg"
    container:
        "docker://faizanbashir/python-datascience:3.6"
    log:
        "logs/plot-hist/{country}.log"
    script:
        "scripts/plot-hist.py"

rule convert_to_pdf:
    input:
        "{prefix}.svg"
    output:
        "{prefix}.pdf"
    log:
        "logs/convert-to-pdf/{prefix}.log"
    wrapper:
        "0.47.0/utils/cairosvg"
```

Legend
- domain knowledge
- technical knowledge
- Snakemake knowledge
- general education

b



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])
```

d