Track patient recovery in real-time by processing streaming data

BIOMEDICAL DATA DESIGN

TA: Haoyin Xu

Group: Zhenyu Xiao
Haobin Zhou
Yimeng Xu
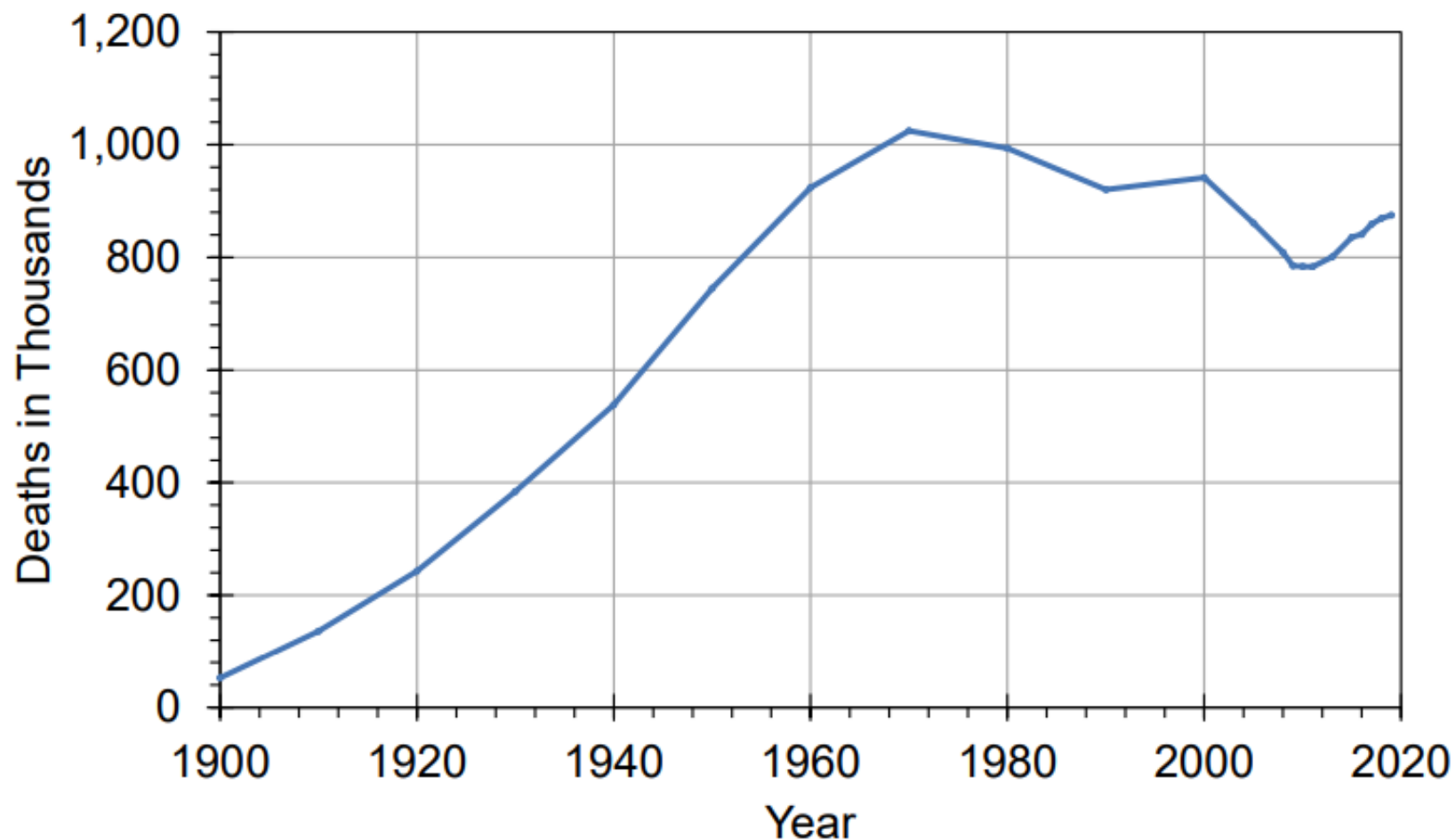Emma Cardenas

Content

01

Goal & Literature Review
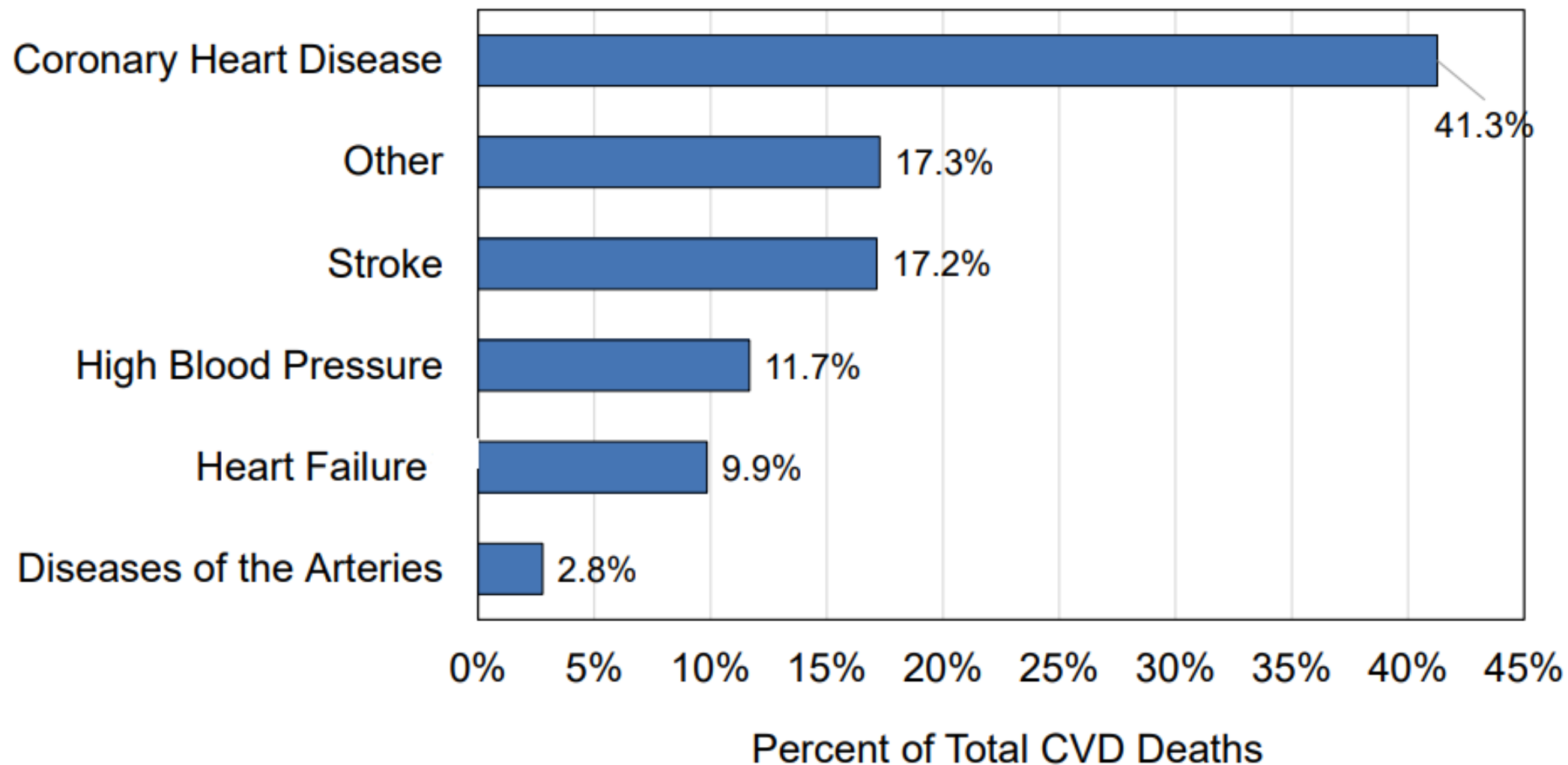
Goal & Literature Review

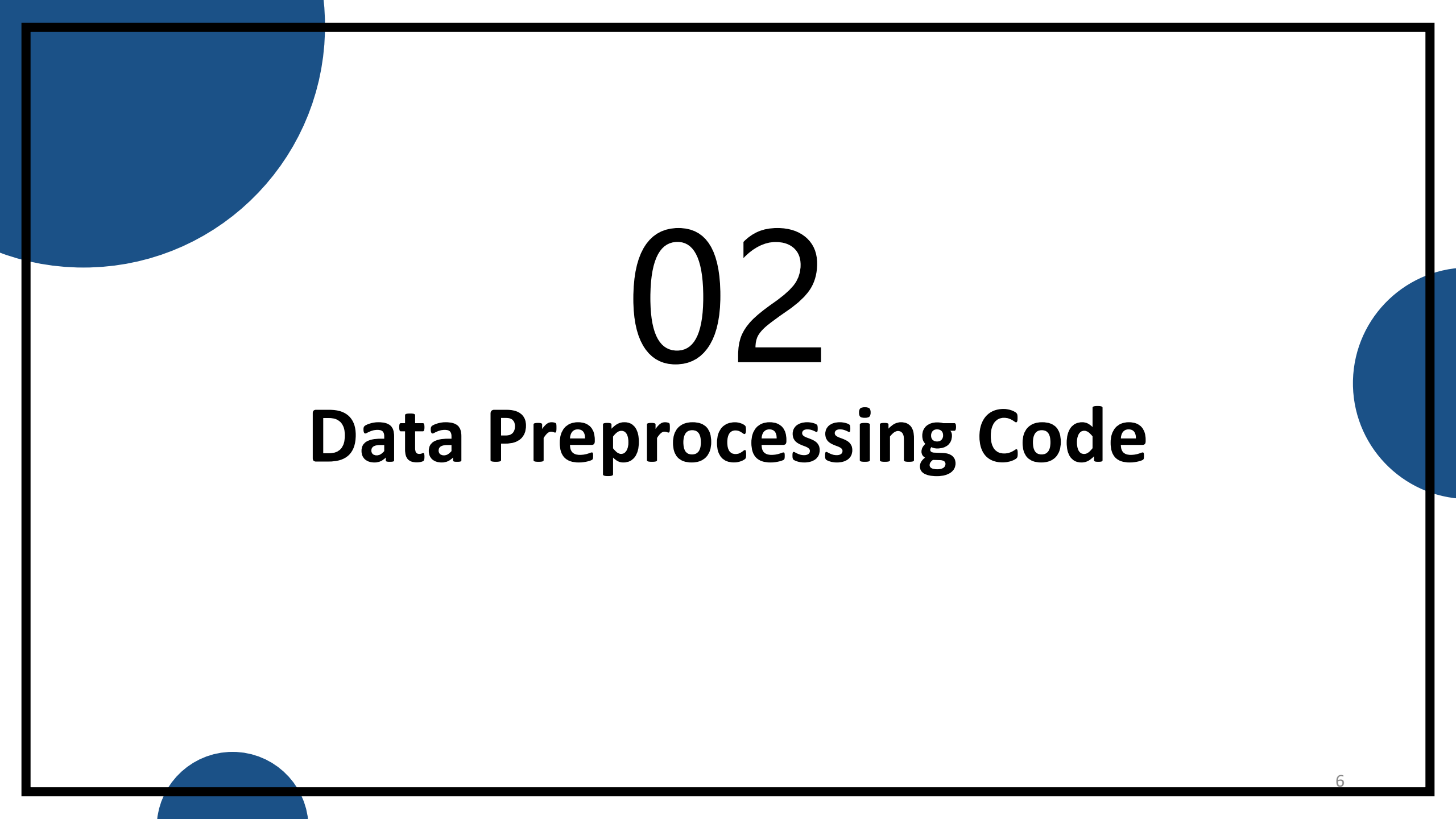## Why we choose cardiovascular diseases?

Deaths attributable to cardiovascular disease, US, 1900-2019.

Goal & Literature Review

Percentage Breakdown of Cardiovascular Disease Deaths in the United States in 2019

02

Data Preprocessing Code

Data Preprocessing Code

	MIMIC III	eICU
Data scale	40,000+	200,000+
Data type	more details like ECG data, medical imaging, surgical information	more monitored information and vital signs
Data Availability and processing difficulty	easier	more complicated

More information on the timeline so the time series information will be more authentic

Data Preprocessing Code

Benchmarking deep learning models on large healthcare datasets

Sanjay Purushotham [a] [1] ✉ , Chuizheng Meng [b] [1] ✉ , Zhengping Che [a] ✉ , Yan Liu [a] 👤 ✉

Show more ∨

```
# # select admissions and all features
createAdmissionList(args)
getItemIdList(args)
filterItemId_input(args)
filterItemId_output(args)
filterItemId_chart(args)
filterItemId_lab(args)
filterItemId_microbio(args)
filterItemId_prescript(args)
processing(args)
collect_mortality_labels(args)
getValidDataset(args)
```

Benchmarking deep learning models on large healthcare datasets

Sanjay Purushotham [a] [1] ✉ , Chuizheng Meng [b] [1] ✉ , Zhengping Che [a] ✉ , Yan Liu [a] 👤 ✉

Show more ⌄

```
# # extract 17 processed features, 17 raw features and 99+ raw features
run_necessary_sqls(args)
get_17_features_processed(args)
get_17_features_raw(args)
get_99plus_features_raw(args)

# # generate time series
get_time_series_sample_17_features_processed(args)
get_time_series_sample_17_features_raw(args)
get_time_series_sample_99plus_features_raw(args)
```

03

Next Step

Thank you