

# Reference Data Sets

# Reference Data Sets

- Open data from HK government  
<https://data.gov.hk>
- Census and Statistics Department, HKSAR  
<http://www.censtatd.gov.hk/home>
- Hong Kong Environmental Statistics  
[http://www.epd.gov.hk/epd/english/resources\\_pub/envir\\_info/envir\\_info.html](http://www.epd.gov.hk/epd/english/resources_pub/envir_info/envir_info.html)

# Reference Data Sets

- Network data (**SNAP**)  
<http://snap.stanford.edu/data/>
- The movie database IMDb  
<http://www.imdb.com/interfaces#plain>
- The World DataBank  
<http://databank.worldbank.org/data/home.aspx>
- Wikipedia dumps  
[http://en.wikipedia.org/wiki/Wikipedia:Database\\_download](http://en.wikipedia.org/wiki/Wikipedia:Database_download)
- Google's Ngram Viewer  
words and phrases (i.e., ngrams) and their usage frequency  
over time of over 8 million books  
<http://storage.googleapis.com/books/ngrams/books/datasetsv2.html>