Przygotowując propozycję mini projektu z HD warto zajrzeć na proponowane niżej strony, na których dostępne są dane źródłowe. Ponadto proszę zwrócić uwagę na ich dostępność, jakość i wielkość zbioru (liczba rekordów reprezentujących fakty oraz przedział czasowy ich wystąpienia – miń. 3 lata).

Poniżej kilkanaście interesujących źródeł, z których możesz pobrać darmowe zbiory danych do eksperymentów z hurtowniami danych:

1. Zbiór datasetów Kaggle.   
   <https://www.kaggle.com/datasets>
2. Baza danych Wide World Importers.  
   [https://docs.microsoft.com/en-us/sql/samples/wide-world-importers-what-is?view=sql-server-ver15](https://docs.microsoft.com/en-us/sql/samples/wide-world-importers-what-is?view=sql-server-ver15&fbclid=IwAR1mVx6Kr3GH0w98ZaCKB1mD8Gqf-YpMbudT0xgNYKAwac5mHYhsNQj9wPE)
3. Baza danych Wiadro danych  
   <https://wiadrodanych.pl/big-data/skad-brac-dane/#Twitter>
4. Zbiór datasetów z Data World.  
    <https://data.world/datasets/open-data?>
5. Wyszukiwanie datasetów od Google.  
   [https://datasetsearch.research.google.com](https://datasetsearch.research.google.com/)
6. Dane udostępnione przez rząd USA.  
   <https://www.data.gov/>
7. Dane z krajów członkowskich Unii Europejskiej.  
   <https://data.europa.eu/euodp/en/data/>
8. Główny Urząd Statystyczny.  
   <https://stat.gov.pl/>
9. Portal dziennikarski (głównie dane z USA).  
   <https://data.fivethirtyeight.com/>
10. GitHub firmy BuzzFeed, ciekawe dane ze świata.  
    <https://github.com/BuzzFeedNews/everything>
11. Program Amazona o dostępie do danych Amazon Web Services.  
    [https://aws.amazon.com/opendata/public-datasets](https://aws.amazon.com/opendata/public-datasets/)
12. Eurostat  
    <https://ec.europa.eu/eurostat/web/main/data/database>