Part 1:

Using Python.

Take the following files, create database structure for them. and load them into PostgreSQL.

Whether you normalize the data is up to you, but both header and record data must be stored.

The structure is documented at

https://www1.ncdc.noaa.gov/pub/data/igra/data/igra2-data-format.txt

https://www1.ncdc.noaa.gov/pub/data/igra/data/data-por/USM00070261-data.txt.zip

https://www1.ncdc.noaa.gov/pub/data/igra/data/data-por/USM00070219-data.txt.zip

https://www1.ncdc.noaa.gov/pub/data/igra/data/data-por/USM00070361-data.txt.zip

https://www1.ncdc.noaa.gov/pub/data/igra/data/data-por/USM00070308-data.txt.zip

https://www1.ncdc.noaa.gov/pub/data/igra/data/data-por/USM00070398-data.txt.zip

Part 2:

Connect to Postgres, digest the data, and write CSVt files. The CSV files should include all data for the weather balloons, and be partitioned by thousands of meters altitude (i.e. partitioning to 0-1000, 1001-2000, 2001-3000, etc). You may create multiple files per partition but each file should only contain data for one partition.

Good to know:

The exercise may require a few tens of GB of space. The exercises can be submitted as a zip archive of the source directory.