**Public data portal examples**

**Gene Expression Omnibus**

<https://www.ncbi.nlm.nih.gov/geo/>

Search query = clear cell renal cell carcinoma

Show additional filters 🡪 sample count

Filters:

Organism = Homo sapiens

Sample count = from 50 to 1000

Study type = Expression profiling by array

Entry type = DataSets (= curated data)

Select the following data: [Clear-cell renal cell carcinoma tumors and tumorgrafts deficient for tumor suppressor BAP1 or PBRM1](https://www.ncbi.nlm.nih.gov/sites/GDSbrowser?acc=GDS4282)

Download: DataSet SOFT file

Inspect file in Sublime Text: show the difference between the annotation lines and the data lines

**Xena data hubs**

<https://xenabrowser.net/hub/>

Select GDC Hub

Select COAD

Download Phenotype + survival data

Inspect files in Sublime text

**ICGC**

<https://dcc.icgc.org/>

Search query = pancreatic cancer

Select project “Pancreatic cancer – CA”: <https://dcc.icgc.org/projects/PACA-CA>

Select “Sequence-based gene expression (EXP-S)”

Filters:

Gender = Female

Download donor data: “Sequencing-based gene expression”

Inspect the file

**UCSC Genome Browser**

<https://genome.ucsc.edu/cgi-bin/hgTracks>

Tools > Table Browser

Filters:

Clade = Mammal

Genome = Human

Assembly = hg38

Group = Regulation

Track = CpG Islands

Output file = cpgi.txt

Inspect the file

**Ensembl BioMart**

<https://www.ensembl.org/biomart/martview>

Choose database Ensembl Genes 97

Choose dataset Human Genes (GRCh38.p12)

Filters:

Region = Chromosome/scaffold 1 to 22

Attributes:

Gene Stable ID

HGNC symbol

Export all results to File TSV Unique results only Go

Inspect the file