


 

☒ Repository
 ☐ Web
 

[View ALL Data Sets](#)

## gene expression cancer RNA-Seq Data Set

*Download:* [Data Folder](#), [Data Set Description](#)

**Abstract:** This collection of data is part of the RNA-Seq (HiSeq) PANCAN data set, it is a random extraction of gene expressions of patients having different types of tumor: BRCA, KIRC, COAD, LUAD and PRAD.

<b>Data Set Characteristics:</b>	Multivariate	<b>Number of Instances:</b>	801	<b>Area:</b>	Life
<b>Attribute Characteristics:</b>	Real	<b>Number of Attributes:</b>	20531	<b>Date Donated</b>	2016-06-09
<b>Associated Tasks:</b>	Classification, Clustering	<b>Missing Values?</b>	N/A	<b>Number of Web Hits:</b>	18645

### Source:

Samuele Fiorini, [samuele.fiorini '@' dibris.unige.it](mailto:samuele.fiorini '@' dibris.unige.it), University of Genoa, redistributed under Creative Commons license (<http://creativecommons.org/licenses/by/3.0/legalcode>) from <https://www.synapse.org/#!Synapse:syn4301332>.

### Data Set Information:

Samples (instances) are stored row-wise. Variables (attributes) of each sample are RNA-Seq gene expression levels measured by illumina HiSeq platform.

### Attribute Information:

A dummy name (gene\_XX) is given to each attribute. Check the original submission ([\[Web Link\]#!Synapse:syn4301332](https://www.synapse.org/#!Synapse:syn4301332)), or the platform specs for the complete list of probes name. The attributes are ordered consistently with the original submission.

## Relevant Papers:

Weinstein, John N., et al. 'The cancer genome atlas pan-cancer analysis project.' Nature genetics 45.10 (2013): 1113-1120.

## Citation Request:

The original data set (hosted at [\[Web Link\]](#)#!Synapse:syn4301332) is maintained by the cancer genome atlas pan-cancer analysis project.

---

Supported By:



In Collaboration With:



[About](#) || [Citation Policy](#) || [Donation Policy](#) || [Contact](#) || [CML](#)