**User Guide for “analyze10xgenomCancer” Function**

**Overview**

The **analyze10xgenomCancer** function is designed to process 10x Genomics single-cell RNA-seq data, specifically aimed at cancer research. This function handles data loading, preprocessing, clustering, and visualization, along with advanced analyses like gene correlation network construction and the calculation of Hausdorff distances for cluster comparisons, and Cancer classification.

**Usage**

The function can be used in a default mode with minimal required inputs but also allows for extensive customization through various parameters.

**Parameters**

* **file\_path**: Path to the file containing the gene expression matrix (HDF5 format).
* **min\_features** (default = 0): Minimum number of features (genes) to filter cells.
* **max\_features** (default = 25000): Maximum number of features (genes) for cell filtering.
* **mt\_threshold** (default = 100): Threshold for mitochondrial content filtering. Cells above this threshold will be excluded.
* **dims** (default = 1:10): Dimensions to use for PCA and UMAP.
* **resolution** (default = 0.2): Clustering resolution for identifying cell populations.

**Example Usage**

Here are two example commands to run the **analyze10xgenomCancer** function based on different user needs:

1. **Basic Usage** (no advanced configurations, just data processing):

analyze10xgenomCancer("file\_path \\V1\_Breast\_Cancer\_Block\_A\_Section\_1\_filtered\_feature\_bc\_matrix.h5")

1. **Advanced Usage** (with custom configurations for detailed analysis):

analyze10xgenomCancer( file\_path = "file\_path \\V1\_Breast\_Cancer\_Block\_A\_Section\_1\_filtered\_feature\_bc\_matrix.h5", min\_features = 200, max\_features = 6000, mt\_threshold = 10, dims = 1:15, resolution = 0.3 )

**Output**

* **SeuratObject**: A Seurat object containing the processed single-cell data.
* **UMAPPlot**: A plot object showing the UMAP visualization of the data.
* **Hausdorff distance matrices:** Distance matrix in different threshold
* **Distribution of Frobenius Norm:** Changes of Frobenius Norm in different threshold.
* **LinePlot**: A plot object showing changes of Frobenius Norm in different threshold.

**Files Generated**

* **SigGenes.txt** and **SigClusters.txt**: These files will contain significant genes and clusters identified during the analysis, formatted based on the specified FDR level in subsequent versions of the function that might handle statistical testing.

**Error Handling**

If there is an issue with the data input or during any of the processing steps, the function will return an error message detailing what went wrong, ensuring that users can make necessary adjustments.