Data Definition for Dataset: kangaroo rat

Dataset Link: https://hgdownload.soe.ucsc.edu/goldenPath/dipOrd1/bigZips/genes/

1. Chromosome: This column represents the scaffold or chromosome number where the genomic feature is located. It helps identify the specific genomic region within the organism's genome.
2. Gene Type: This column specifies the type of gene. It provides information about the classification or source of the gene, which can help understand its origin or category (e.g., ensGene).
3. Feature Type: This column indicates the type of genomic feature being described. It can represent various components of the gene, such as transcripts or exons. Common values in this column may include "transcript" or "exon."
4. Start Position: This column denotes the starting position of the genomic feature on the chromosome. It provides the genomic coordinate where the feature begins, typically represented as a numerical value.
5. End Position: This column indicates the ending position of the genomic feature on the chromosome. It provides the genomic coordinate where the feature ends, typically represented as a numerical value.
6. Strand: This column represents the orientation of the genomic feature on the chromosome. The strand can be either "+" or "-", denoting the forward or reverse orientation, respectively.
7. Gene ID: This column provides a unique identifier for each gene. It assigns a specific identifier to each gene in the dataset, facilitating easy identification and retrieval of gene-related information.
8. Transcript ID: This column offers a unique identifier for each transcript associated with the gene. It allows for specific identification and tracking of individual transcripts, which are RNA molecules transcribed from the gene.
9. Exon Number: This column specifies the number of the exon within the transcript, if applicable. Exons are the coding regions of a gene that are transcribed into RNA and eventually translated into protein.
10. Exon ID: This column provides a unique identifier for each exon, if applicable. It assigns a specific identifier to each exon, aiding in the identification and analysis of individual exons within the dataset.
11. Gene Name: This column represents the name or identifier assigned to the gene. It provides a more human-readable identifier or name for the gene, which can be helpful for referencing and understanding the gene's identity.

Data Definition for Dataset: Rat

Dataset Link: <https://hgdownload.soe.ucsc.edu/goldenPath/rn7/bigZips/genes/>

1. Chromosome: This column represents the chromosome number or identifier where the gene is located. It provides information about the specific chromosome in the rat genome.
2. Gene Type: This column specifies the type or source of the gene data. It indicates the database or reference from which the gene information is obtained, such as ncbiRefSeq.
3. Feature Type: This column indicates the type of genomic feature being described. It can represent various components of the gene, such as transcripts or exons. Common values in this column may include "transcript" or "exon".
4. Start Position: This column denotes the starting position of the genomic feature on the chromosome. It provides the genomic coordinate where the feature begins, typically represented as a numerical value.
5. End Position: This column indicates the ending position of the genomic feature on the chromosome. It provides the genomic coordinate where the feature ends, typically represented as a numerical value.
6. Strand: This column represents the orientation of the genomic feature on the chromosome. The strand can be either "+" or "-", denoting the forward or reverse orientation, respectively.
7. Gene ID: This column provides a unique identifier for each gene. It assigns a specific identifier to each gene in the dataset, facilitating easy identification and retrieval of gene-related information.
8. Transcript ID: This column offers a unique identifier for each transcript associated with the gene. It allows for specific identification and tracking of individual transcripts, which are RNA molecules transcribed from the gene.
9. Exon Number: This column specifies the number of the exon within the transcript, if applicable. Exons are the coding regions of a gene that are transcribed into RNA and eventually translated into protein.
10. Exon ID: This column provides a unique identifier for each exon, if applicable. It assigns a specific identifier to each exon, aiding in the identification and analysis of individual exons within the dataset.
11. Gene Name: This column represents the name or identifier assigned to the gene. It provides a more human-readable identifier or name for the gene, which can be helpful for referencing and understanding the gene's identity.