R: An Introduction - Homework

Homework 1

Before you start, create a project folder that includes the brauer2007_tidy.csv dataset and your r script.

EXERCISE 1

R as a Calculator

I want to calculate how fast I drove from one location to another, however, for some odd reason I only know how many kilometers I traveled and the number of minutes it took me. I really only understand speed when it comes at me in the form of miles per hour. Can you show me how fast I traveled in miles per hour? Was I driving too fast?

1. Create a variable called `kilometers` and set it equal to 120
2. Create a variable called `minutes` and set it equal to 40
3. Convert `minutes` to `hours` by dividing `minutes` by 60
4. Convert `kilometers` to `miles` by multiplying `kilometers` by .62
5. Calculate the `mph` (miles per hour)

EXERCISE 2

Calculate the cosine of the square root of the absolute value of (2^(25-21) - 40). `2`.

EXERCISE 3

Working with vectors

1. Create a vector that starts at 3 and goes to 300, increasing by 7 at each step.
2. Multiply the above vector by 3 and save it as a new vector
3. Calculate the square root of the vector created in question 2. Store this object in a new vector.
4. Return the first 5 elements of the vector from the object created in question 3.
5. Return the result of elements 3, 5, and 8 multiplied by elements 7, 15, and 23 from the vector created in question 3.

EXERCISE 4

Working with real data

The data we're going to look at is cleaned up version of a gene expression dataset from Brauer et al. Coordination of Growth Rate, Cell Cycle, Stress Response, and Metabolic Activity in Yeast (2008) *Mol Biol Cell* 19:352-367. This data is from a gene expression microarray, and in this paper the authors are examining the relationship between growth rate and gene expression in yeast cultures limited by one of six different nutrients (glucose, leucine, ammonium, sulfate, phosphate, uracil). If you give yeast a rich media loaded with nutrients except restrict the supply of a *single* nutrient, you can control the growth rate to any rate you choose. By starving yeast of specific nutrients you can find genes that:

Raise or lower their expression in response to growth rate. Growth-rate dependent expression patterns can tell us a lot about cell cycle control, and how the cell responds to stress. The authors found that expression of >25% of all yeast genes is linearly correlated with growth rate, independent of the limiting

nutrient. They also found that the subset of negatively growth-correlated genes is enriched for peroxisomal functions, and positively correlated genes mainly encode ribosomal functions.

Respond differently when different nutrients are being limited. If you see particular genes that respond very differently when a nutrient is sharply restricted, these genes might be involved in the transport or metabolism of that specific nutrient.

1. Open up the `brauer2007_tidy.csv` data and inspect the data.
2. Import the data into R, saving it as `brauer`.
3. What's the range of `expression` values represented in the data?
4. What class of variable are the `bp` and `mf` columns?
5. What is the mean `expression` value?
6. How many times does Uracil show up in the `nutrient` column?
7. How many missing values are there in the `mf` column?