# Homework 1

AUTHOR Aysha Talakkal Cosumal - 23200212 PUBLISHED
June 17, 2024

#### INTRODUCTION

In this document, we will explore the **Orange** dataset, which is a small dataset with 35 recorded observations and 3 columns, recording the growth of orange trees.

## **DATASET DESCRIPTION**

Let us look at the first few rows of our dataset and a detailed description of the variables:

	Tree	age	circumference	
1	1	118	30	
2	1	484	58	
3	1	664	87	
4	1	1004	115	
5	1	1231	120	
6	1	1372	142	
7	1	1582	145	
8	2	118	33	
9	2	484	69	
10	2	664	111	

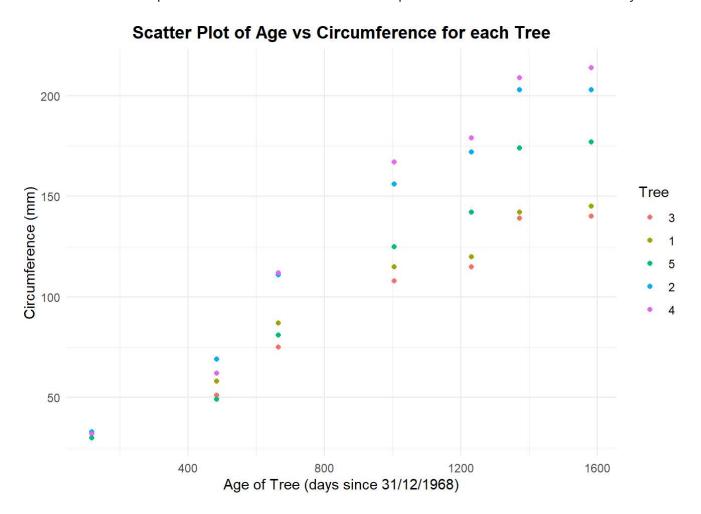
- *Tree*: ordered factor, indicates which tree measurement is being made on
- *age*: a numeric vector giving the age of the tree (days since 31st December, 1968)
- *circumference*: a numeric vector of trunk circumferences (mm)

Let us look at a summary of our variables, age and circumference:

age			circumference				
	Min.	:	118.0	Min.	:	30.	0
	1st Qu.	:	484.0	1st Qu.	:	65.	5
	Median	:1	L004.0	Median	:1	15.	0
	Mean	:	922.1	Mean	:1	15.	9
	3rd Qu.	:1	L372.0	3rd Qu.	:1	61.	5
	Max.	:1	L582.0	Max.	: 2	14.	0

#### **SCATTERPLOT**

Let us create a scatterplot of our data to see the relationship between our variables more clearly:



## **DESCRIPTION**

The above scatterplot provides a visual representation of the relationship between age of the tree in days (since 31st December, 1968) and its trunk circumference. This plot uses the color argument to differentiate between the trees, so that it is easier to infer from.

From the plot, we can clearly see that there is a positive relationship between Age and Circumference. This means that as the tree grows older, the trunk of the tree grows thicker. These findings are consistent with the known fact, that knowing the girth of a tree can tell us its age.