In this worksheet we will revisit the birthday problem. That is, we have N people in a room. We further assume that all years have 365 days and that birthrates are constant throughout[1]. One key question of interest is to find the

$$\begin{aligned} &P(\text{at least two people in the room who share a birthday}) \\ &= 1 - P(\text{no two people have the same birthday}) \\ &= 1 - \frac{365}{365} \cdot \frac{364}{365} \cdots \frac{365-(N-1)}{365}. \end{aligned}$$

We will now compute the above probabilities in R as N varies and compare the values.

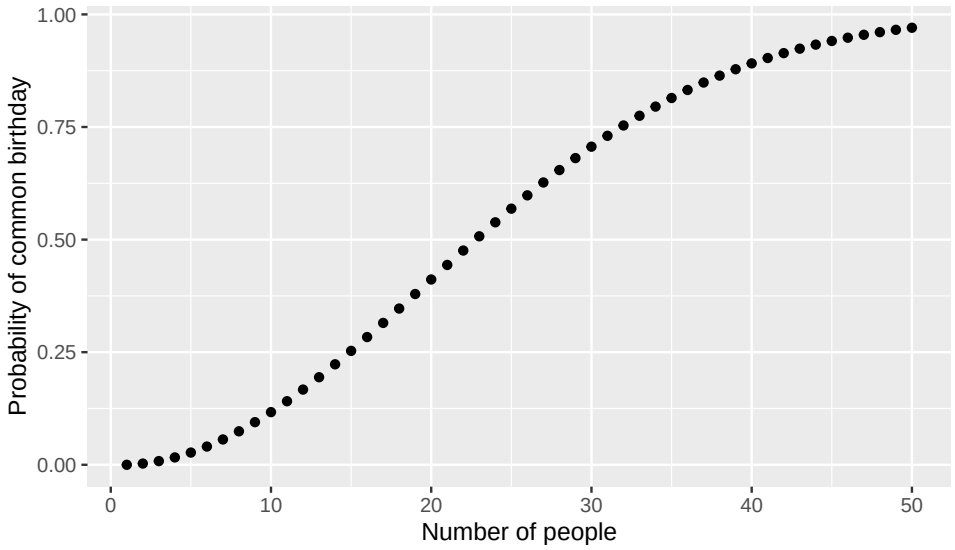
1. Let's take $N = 10$. The below R-code computes above probability when there are 10 people in the room.

```
> N=10
> 1-(choose(365,N)*factorial(N))/365^N
```

   Use the `help(factorial)` in R and verify for yourself that the code is indeed computing the above probability and is providing a numerical answer[2] equal to $1-\prod_{i=0}^{9} \frac{365-i}{365}$.

2. The below R-code plots the above probability as a function of N.

```
> library(ggplot2)
> num=1:50
> prob=c()
> for(i in 1:50){prob[i]=1-(choose(365,i)*factorial(i))/365^i}
> ggplot(data.frame(num,prob))+geom_point(aes(x=num,y=prob))+
+   xlab('Number of people')+ylab('Probability of common birthday')
```



[1] i.e. it is assumed that for each personl all 365 possible birthdays are equally likely.

[2] Another interesting exercise in R would be to use the `for(i in 0:9){ ... }` loop construct to do the above computation.

The above is an implementation in ggplot which is a system for declaratively creating graphics, created by Hadley Wickham in 2005, -based on Leland Wilkinson's Grammar of Graphics. To understand the syntax and the methodology we refer the reader to https://ggplot2.tidyverse.org/.

(a) What is the value of N above which there is a 60% chance that two of the N people will have a common birthday ?

(b) If $N = 20$ people in a room, from the plot, infer what is chance that two of the 20 people will have a common birthday ?

(c) In the same room of N people, what are the chances that 3 people share the same birthday ? Can you a write a R-code to implement the same for k people ?

3. To save the plot as a jpeg file:

```
> jpeg(filename="Birthday.jpeg", width=6, height=6, units="in", res=500)
> ggplot(data.frame(num,prob))+geom_point(aes(x=num,y=prob))+
+   xlab('Number of people')+ylab('Probability of common birthday')
> dev.off()
```