Chapter 18: Sampling Function and Simulation

# Learning Objectives

After reading this chapter, you will be able to

* Distinguish between population and sample in a business reality.
* Understand sample function in   
  excel
* Understand the use of RAND function
* Analyse problems using Monte-Carlo simulation

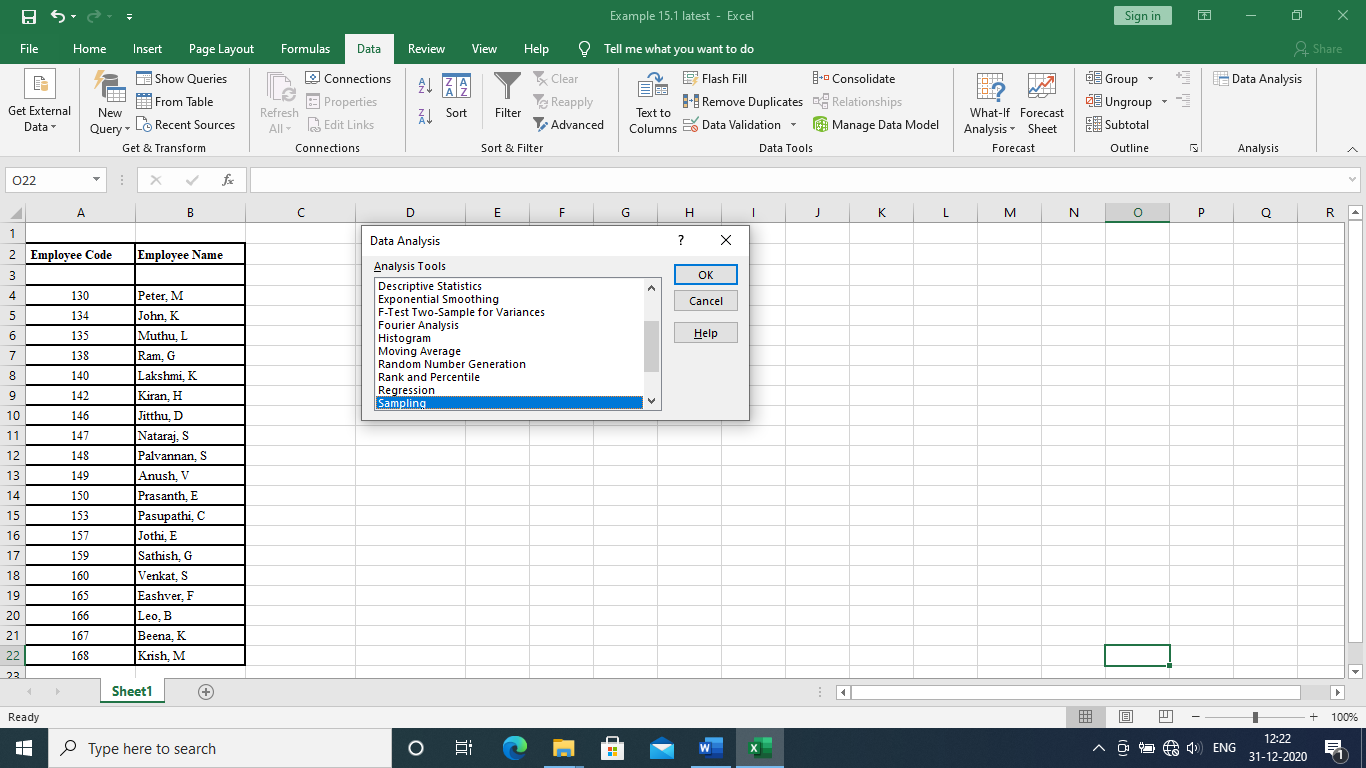
# 18.1. Introduction

A real-life entity with all its members is called a population. A subset of that population for studying about the population is called a sample. All the employees in an organization can form an example of a population. A subset of those employees, say 20 per cent of the employees, from each section of the organization put together will form a sample. The outcome of the study based on the sample can be extended to the population. This chapter presents the capabilities of *sample function* in Excel. Further, RAND function and Monte-Carlo simulation using RAND function are also demonstrated.

# 18.2. Sample Function in Excel

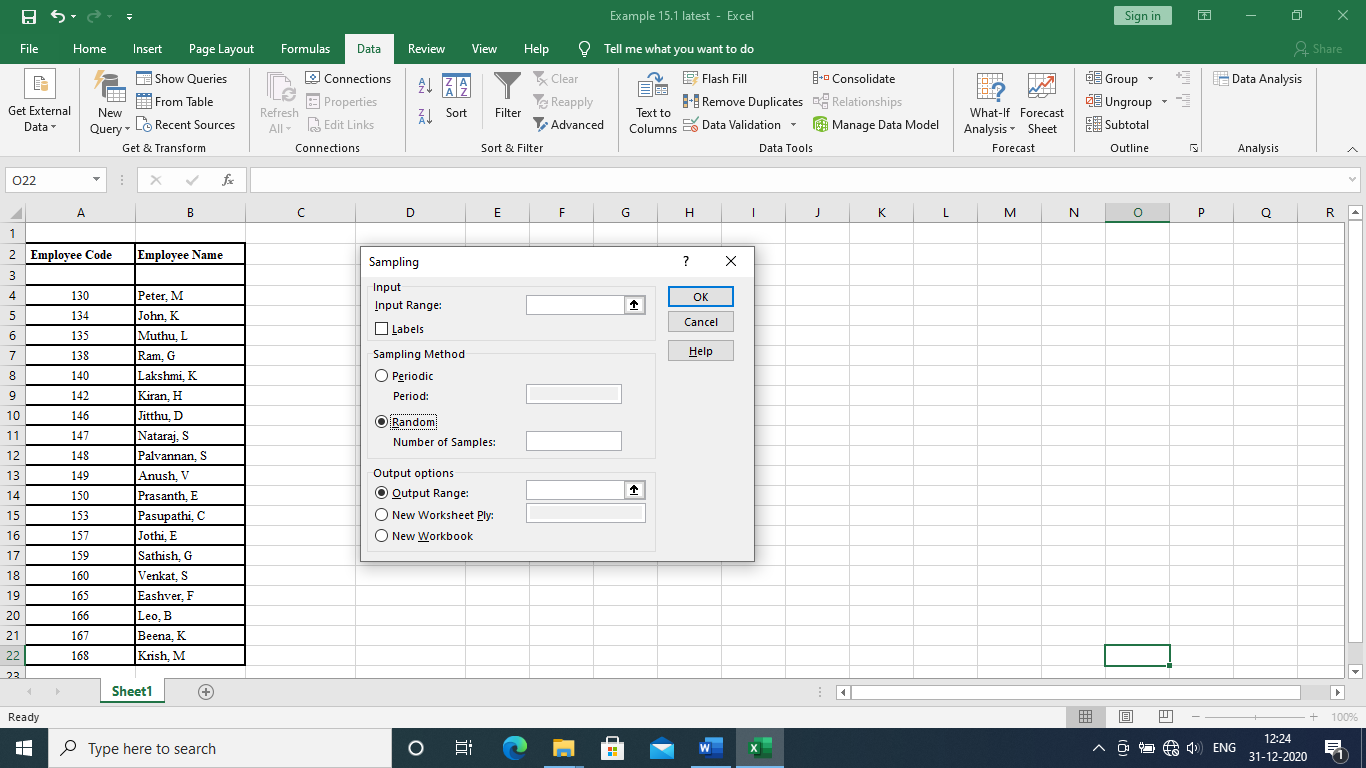
In reality, investigators may be selecting a certain number of sample units from a population. Consider the case of selecting a specified number of respondents (sample units) for a survey from among the employees (population) of a particular cadre in a company. The code number of the employees may be used as a key to select the specified number of employees of that cadre. The sample function can be used to select a desired number of code numbers of the employees from among the code numbers of the employees at a particular cadre. The method of selecting the employees may be either periodic or random. Normally, the random selection method is used. The sampling function in Excel has both options.

The clicks of buttons, "Data → Data Analysis", give a display as shown in Figure 18.1.



**Figure 18.1.** Screenshot of Display of Clicks of Buttons, "Data → Data Analysis"

Now, the click of the button "sampling" in the drop-down menu of Figure 18.1 and click of OK in the same drop-down menu gives a display as shown in Figure 18.2. Now, the investigator has to enter the range of cells containing the data of the population in the box against input range, click labels and then click random. Also, enter the number of samples in the box against the number of samples and then enter the range of cells in the output range to display the selected sample units from the population.



**Figure 18.2.** Screenshot of Display after Clicking Sampling and OK Buttons in the Drop-down Menu of Figure 18.1

#### Example 18.1

An investigator has to select 8 employees from among 18 employees working in a section of a company. The employee codes and their names are shown in Table 18.1. Select eight employees randomly from among 18 employees using sampling function in Excel.

**Table 18.1.** Employee Codes and Names

|  |  |
| --- | --- |
| Employee Code | Employee Name |
| 130 | Peter, M |
| 134 | John, K |
| 135 | Muthu, L |
| 138 | Ram, G |
| 140 | Lakshmi, K |
| 142 | Kiran, H |
| 146 | Jitthu, D |
| 147 | Nataraj, S |
| 148 | Palvannan, S |
| 149 | Anush, V |
| 150 | Prasanth, E |
| 153 | Pasupathi, C |
| 157 | Jothi, E |
| 159 | Sathish, G |
| 160 | Venkat, S |
| 165 | Eashver, F |
| 166 | Leo, B |
| 167 | Beena, K |

#### Solution

The input of the given data of Example 18.1 is shown in Figure 18.3. The display after clicking the buttons, "Data → Data Analysis → Sampling" is shown in Figure 18.4. Then the display after entering "input range", clicking random option, number of samples as 12 and "output range" is shown in Figure 18.5.

*One should note the fact that the number of samples is entered as 12 in place of 8 as given in the problem, because there may be repeating employee codes. At the end, from among the sampled 12 employee codes, non-repeating 8 employee codes from the top of the column may be selected for implementation.*

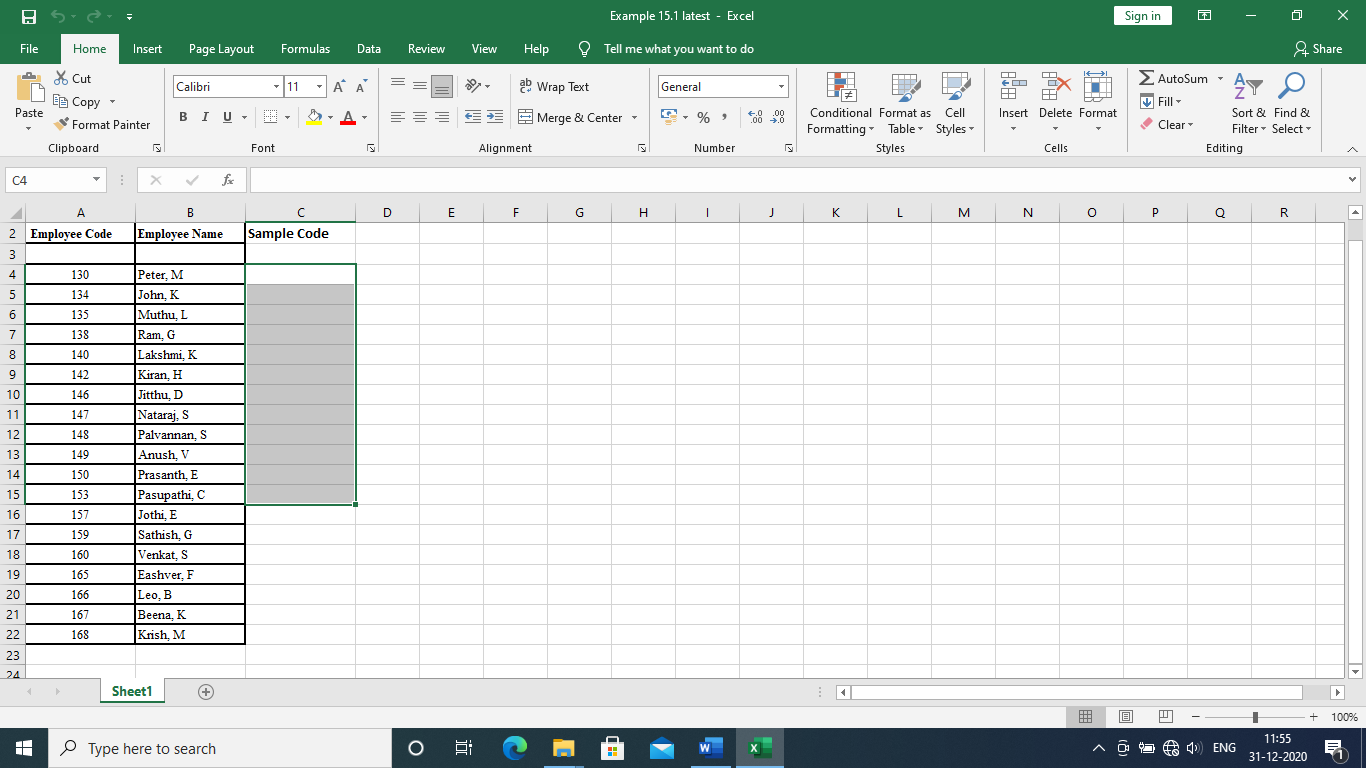
Now, the click of OK button in the drop-down menu of Figure 18.5 gives a display as shown in Figure 18.6.

*An important note is that if the exercise of generating employee codes is repeated using the sampling function with the random option, the entries in the range of cells from C4 to C15 in Figure 18.6 will differ, because the sampling function with the random option uses simulation to generate the employee codes.*

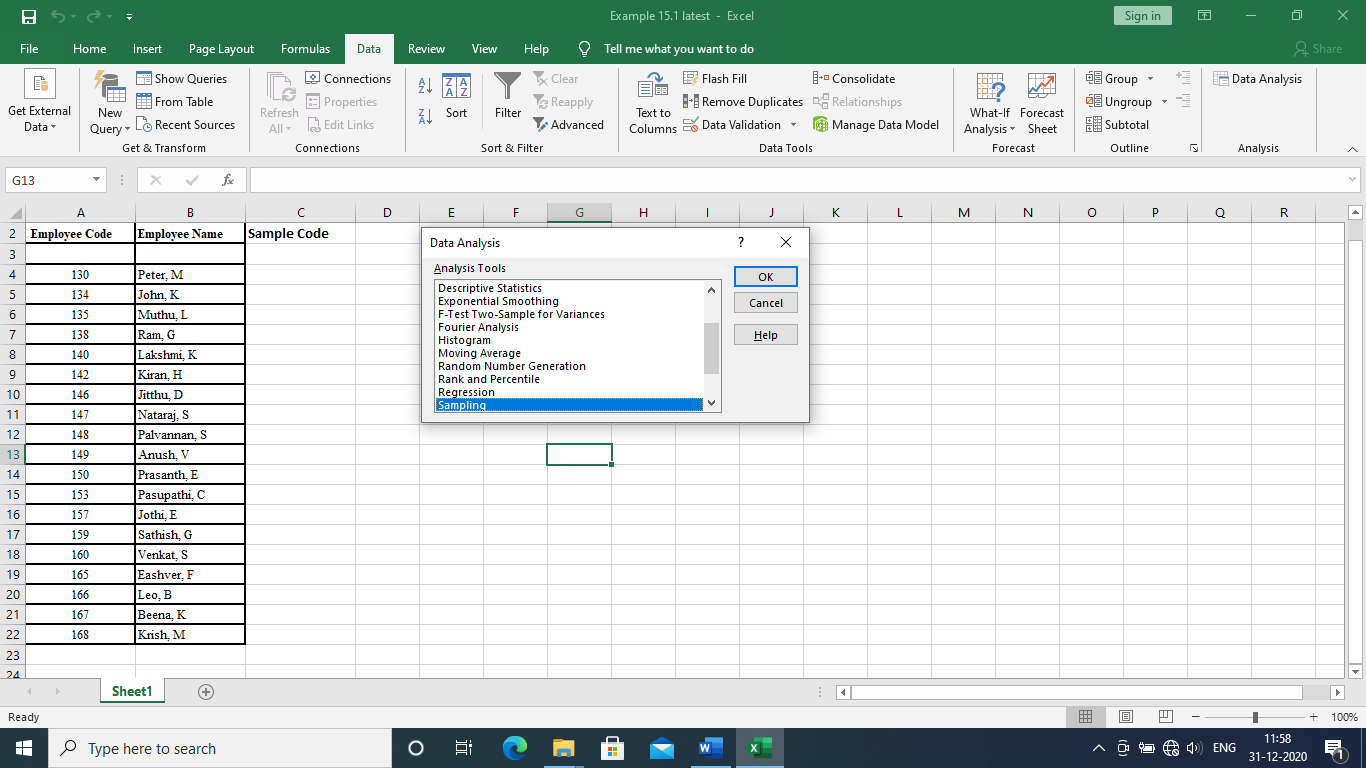
Then select the range C4 to C15 and then select the button "remove duplicates" in the ribbon of Figure 18.6 to have the values in that range without duplicates, which are shown in Figure 18.7. One can verify that three duplicates are removed. Then the eight codes from cell C4 to cell 11 are copied to cell D4 to cell D11, respectively. The serial numbers of these final codes are given in the cells from cells E4 to E11. All these can be seen in Figure 18.7.

For the same problem, the display after clicking periodic sampling method with a period of 2 with other data remain the same as shown in Figure 18.8. Then, clicking OK button in the drop-down menu of Figure 18.8 gives the sampled codes in column C as in Figure 18.9. These sampled codes will not have duplicates because of the periodic sampling method. Then the codes from cell C4 to cell C11 are copied to cell D4 to cell D11, respectively. The serial numbers of these final codes are given from cells E4 to E11. These are shown in Figure 18.10 based on Figure 18.9.

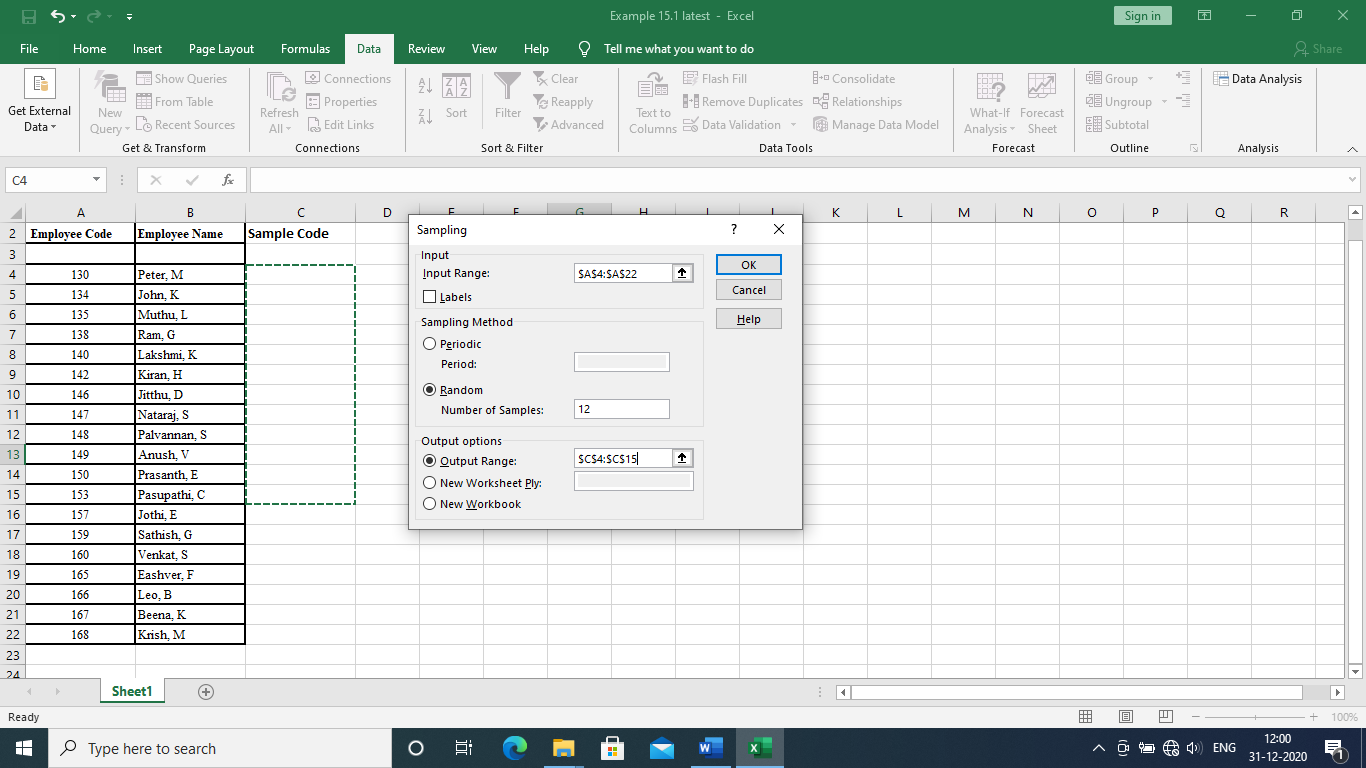
*In both methods, finally manual intervention is required to pick the required number of employee codes.*



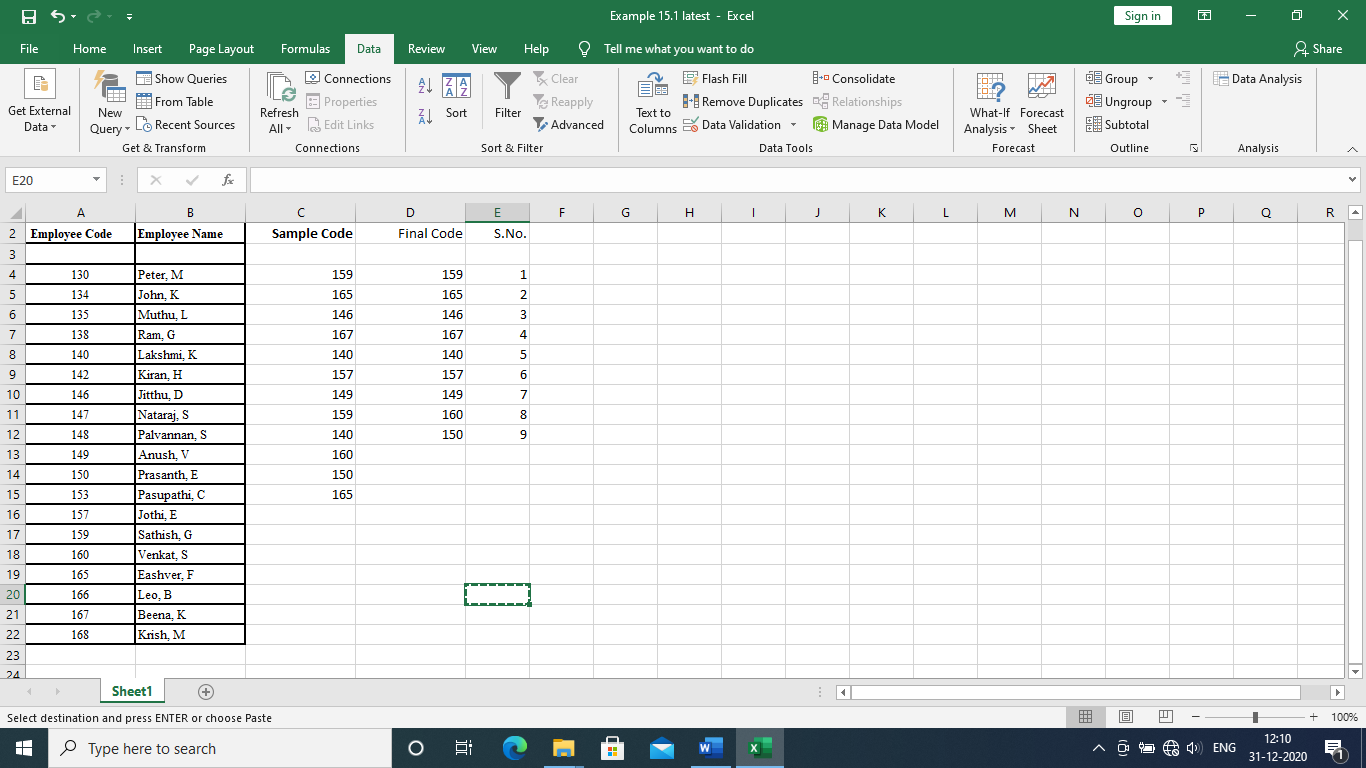
**Figure 18.3.** Screenshot of Input of Example 18.1



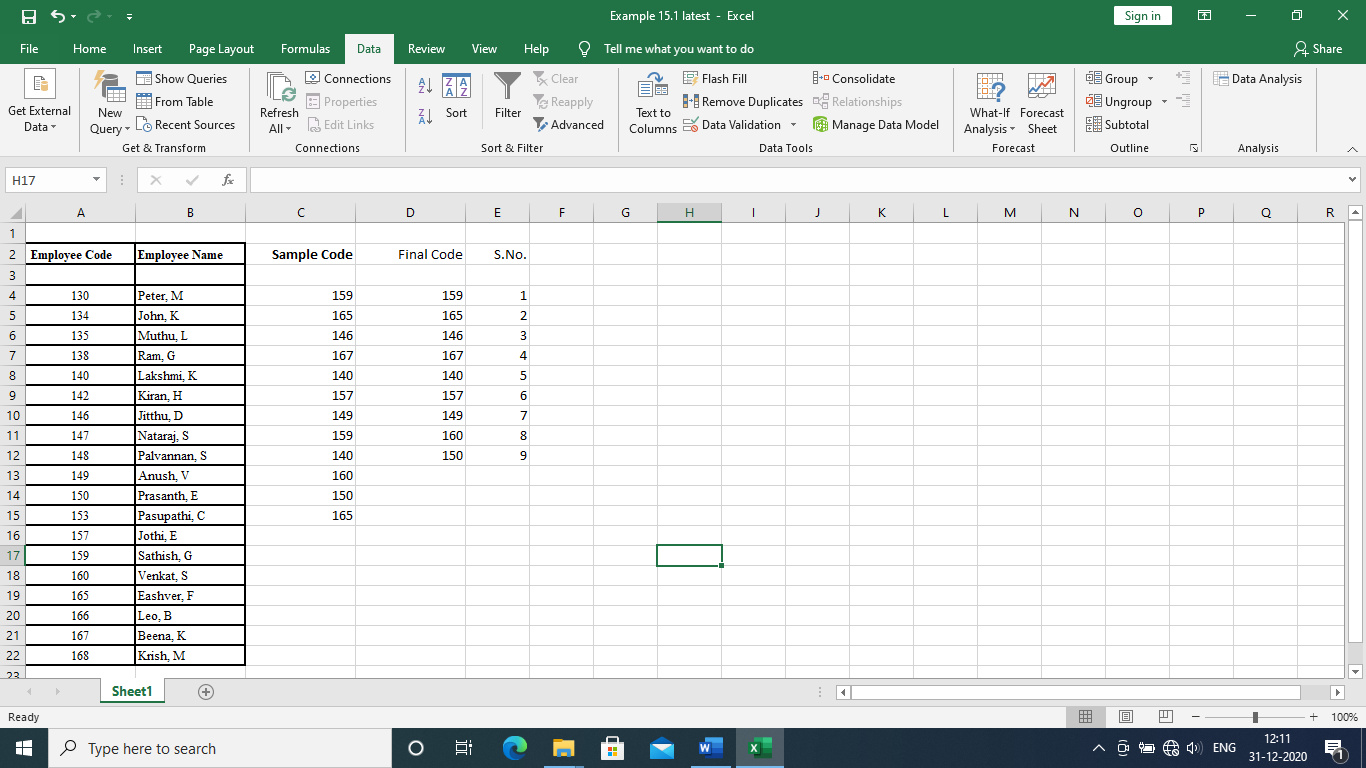
**Figure 18.4.** Screenshot of Display after Clicking Buttons, "Data → Data Analysis → Sampling"



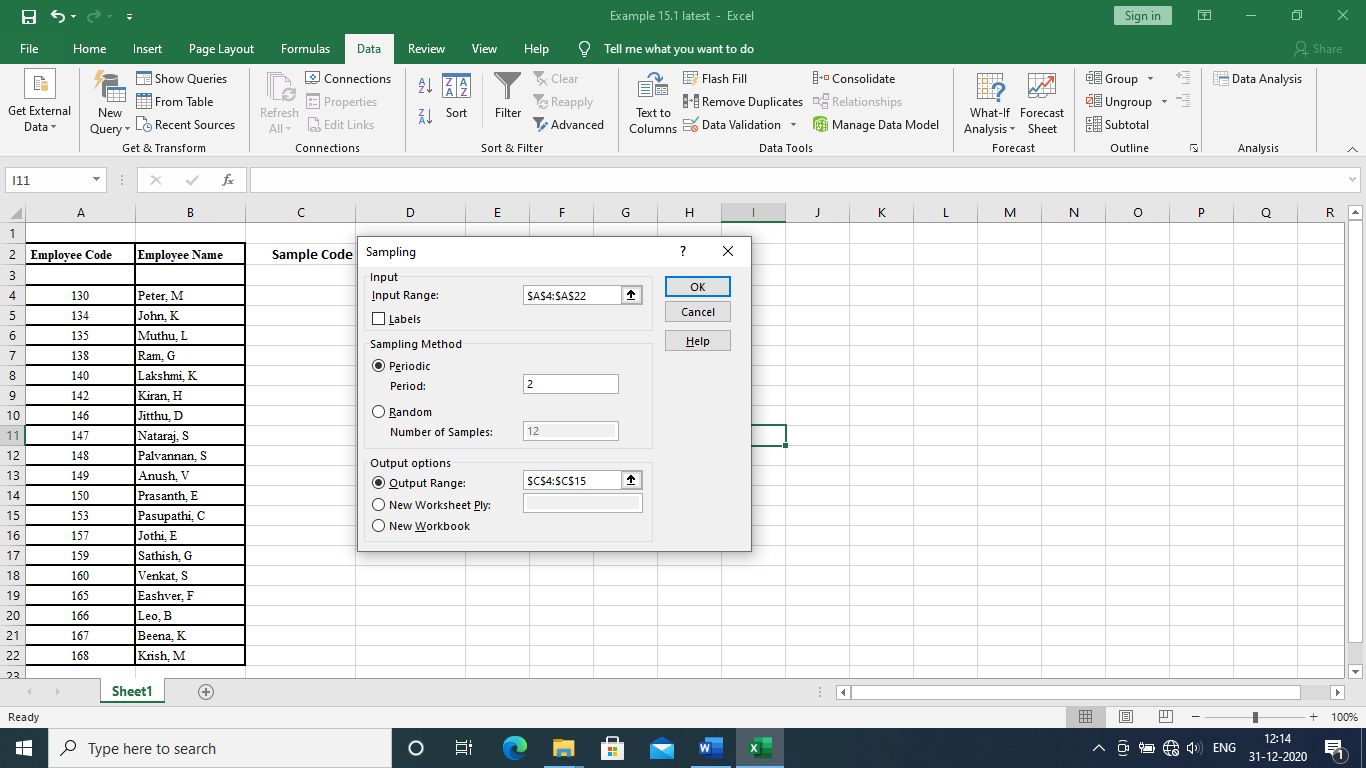
**Figure 18.5.** Screenshot after Entering Needy Details in the Data Boxes in the Drop-down Menu of Figure 18.4



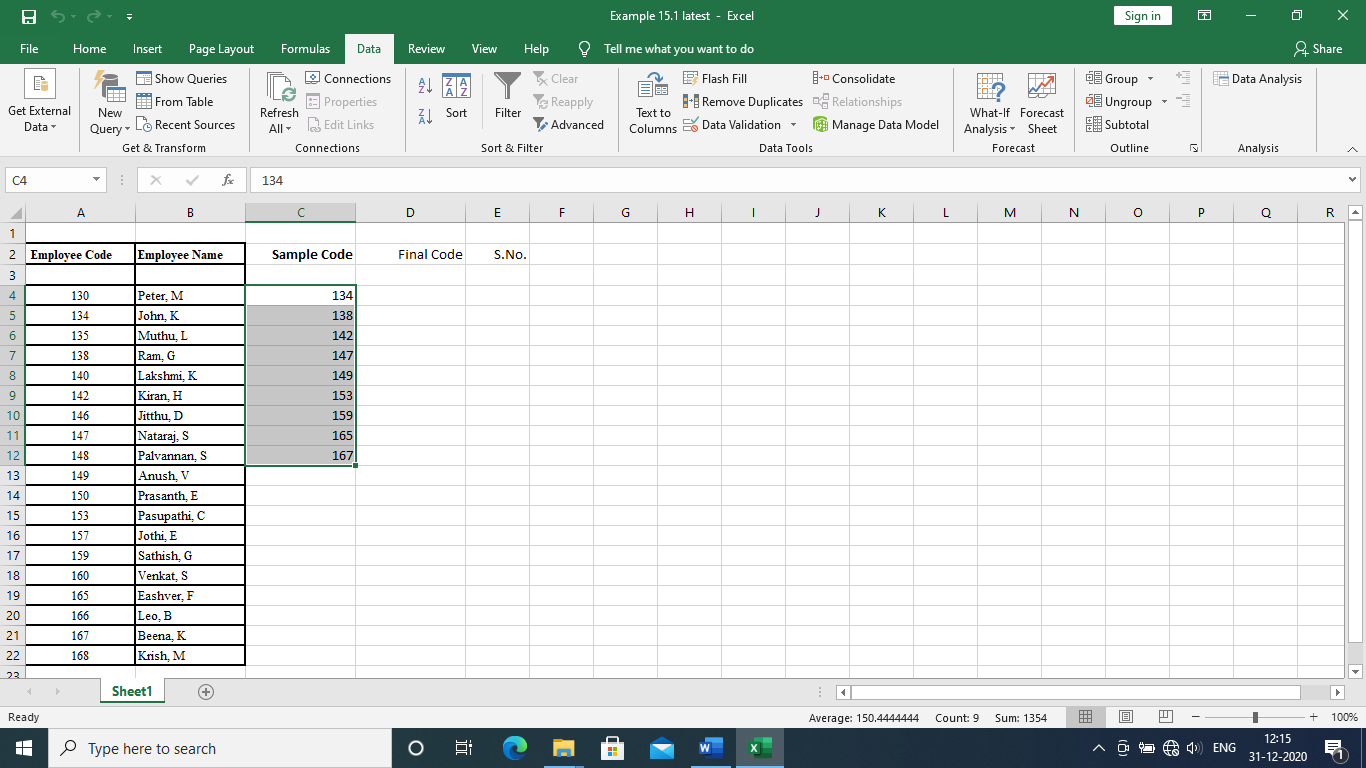
**Figure 18.6.** Screenshot after Clicking OK Button in the Drop-down Menu of Figure 18.5



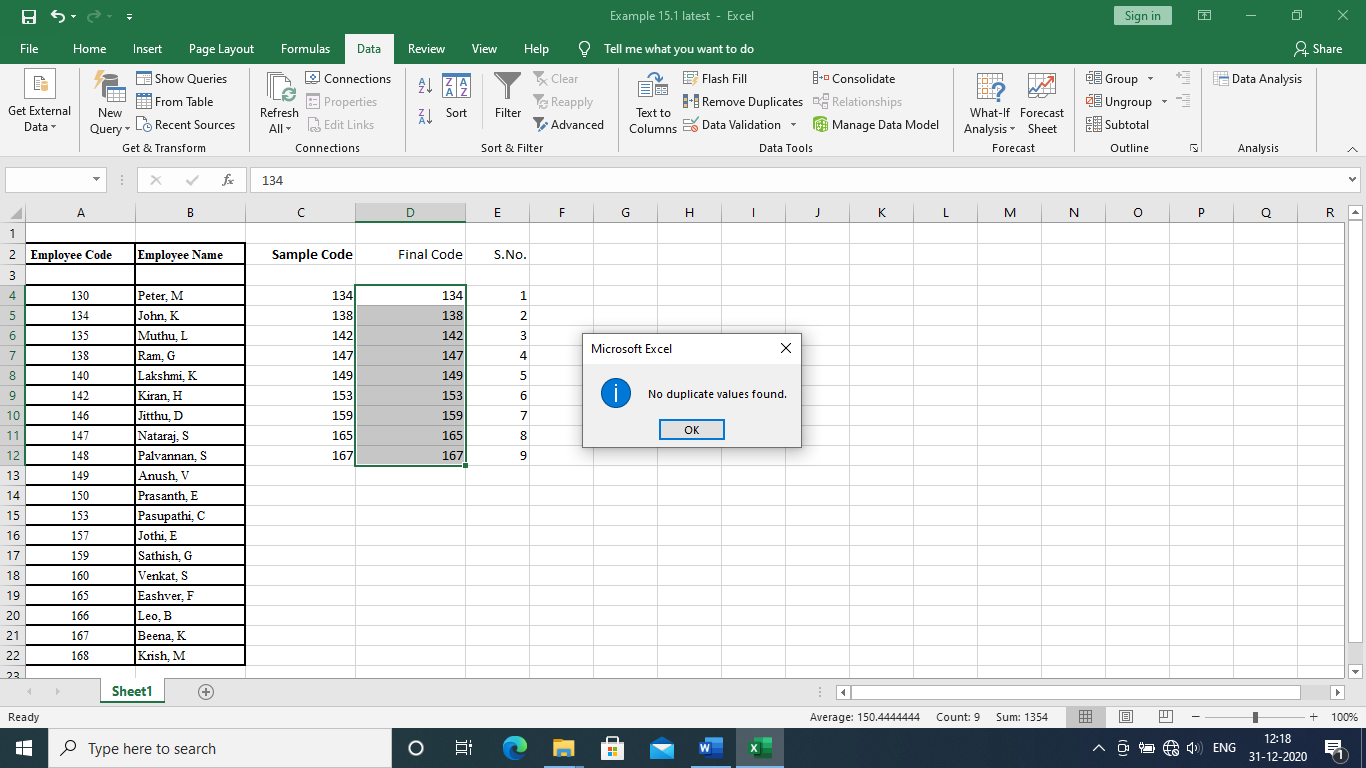
**Figure 18.7.** Screenshot after Clicking "Remove Duplicates" in the Ribbon of Figure 18.6, Copying Eight Sampled Codes to Column D from Column C and Assigning Serial Numbers from 1 to 8 in Column E for Them



**Figure 18.8.** Screenshot of Display for the Option of Periodic Sampling Method with Period 2 along with Data Ranges for Input Range and Output Range



**Figure 18.9.** Screenshot after Clicking OK Button in the Drop-down Menu of Figure 18.8



**Figure 18.10.** Screenshot after Copying Eight Codes from Column C to Column D and Assigning Serial Numbers for Them in Column E in Figure 18.9

# 18.3. Random Number Generation

Random numbers are generated using the following functions.

1. RAND() under Formulas → Math & Trig
2. Random number generation under data analysis under data

## 18.3.1. RAND( ) Function

In many realities, the analyst may require uniformly distributed random numbers, which will be used in the random variate formula of a probability distribution to estimate the value of the respective random variable. The uniformly distributed random number can be obtained using RAND function in Excel.

This function is available after clicking the sequence of buttons, namely Formulas and Math & Trig buttons. Before getting the random number, position the cursor in a desired cell say A2 and then click the sequence of buttons, namely Formulas => Math & Trig => RAND( ), which will give a display as shown in Figure 18.11. Then the click of OK button in the drop-down menu of Figure 18.11 will give a random number in cell A2 as shown in Figure 18.12. This function does not have any argument.

The same can be obtained by using the following formula after positioning the cursor in   
cell A2.

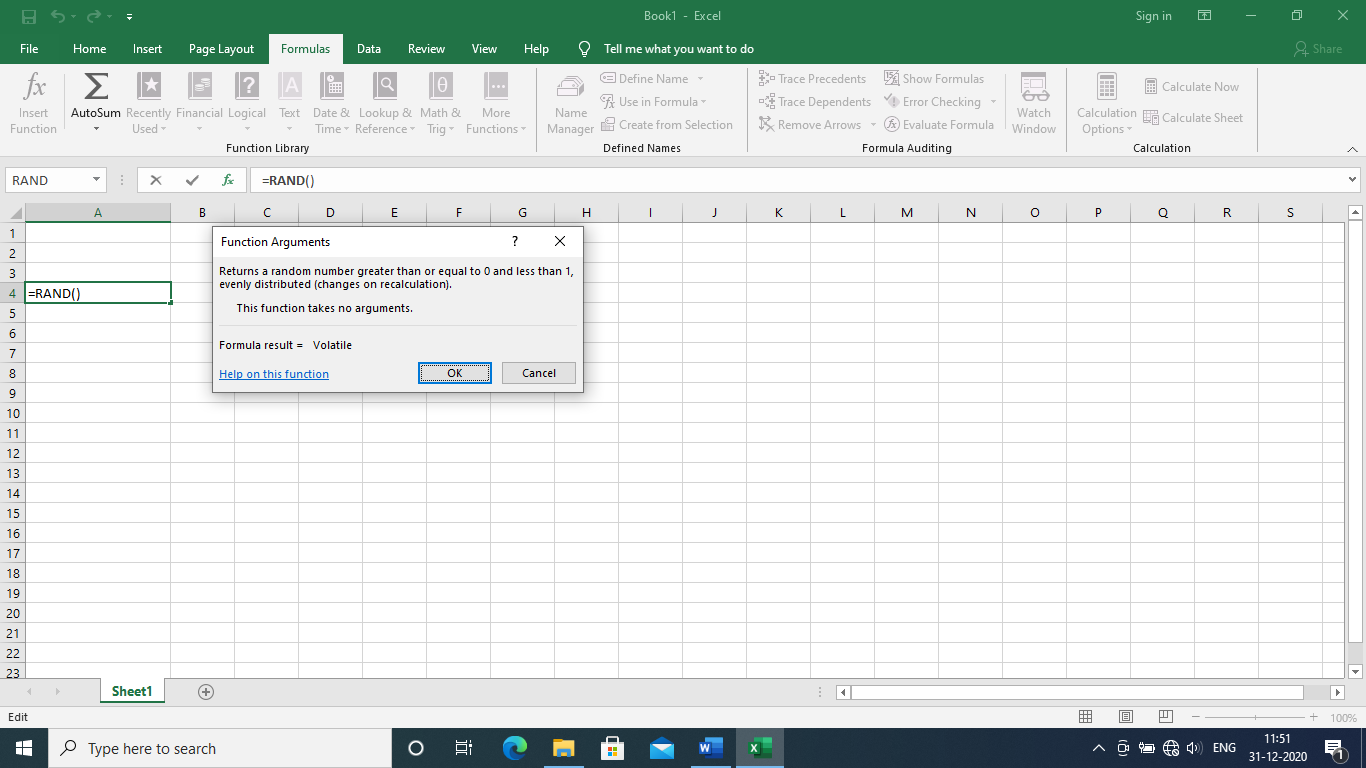
Video provides a powerful way to help you prove your point. When you click online video, you can paste in the embed code for the video you want to add. You can also type a keyword to search online for the video that best fits your document.

To make your document look professionally produced, Word provides header, footer, cover page and textbox designs that complement each other. For example, you can add a matching cover page, header and sidebar. Click insert and then choose the elements you want from the different galleries.

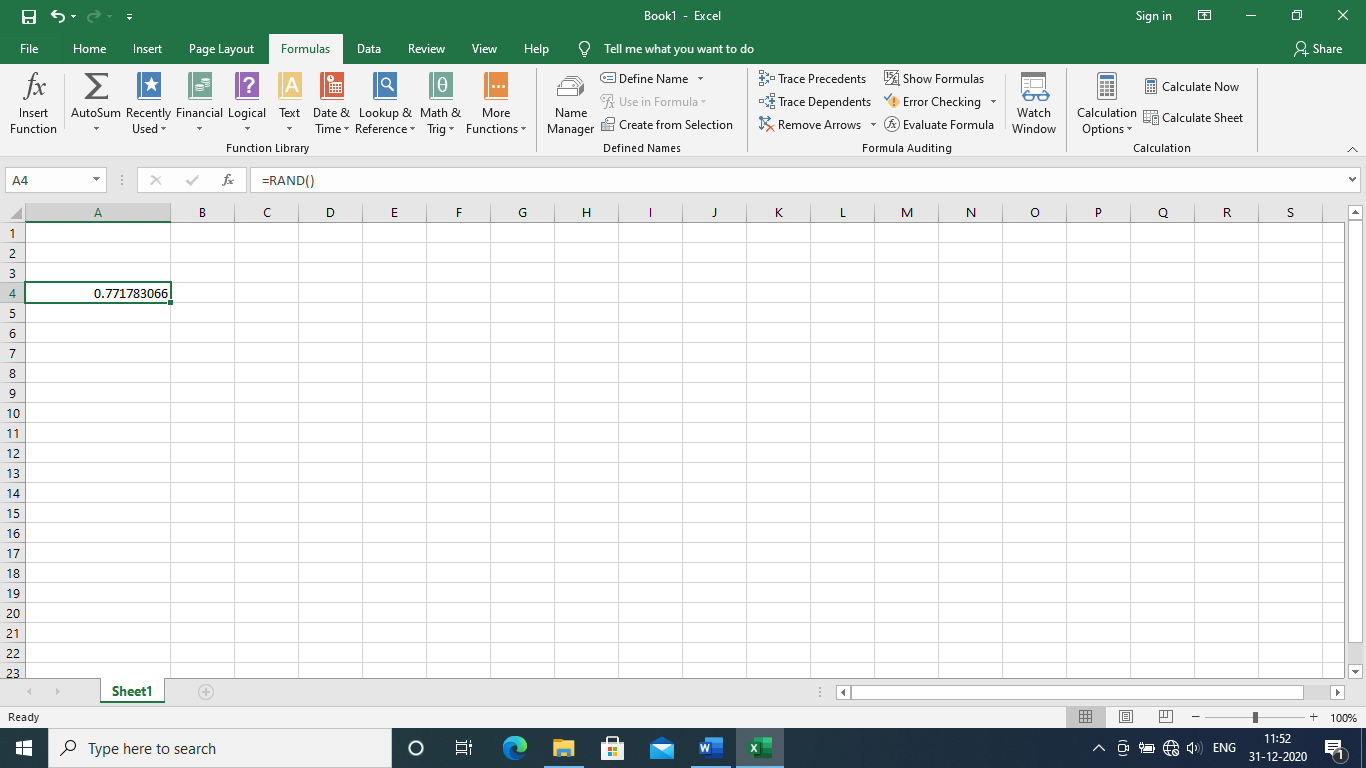
Themes and styles also help keep your document coordinated. When you click design and choose a new theme, the pictures, charts and SmartArt graphics change to match your new theme. When you apply styles, your headings change to match the new theme.

Save time in Word with new buttons that show up where you need them. To change the way a picture fits in your document, click it and a button for layout options appears next to it. When you work on a table, click where you want to add a row or a column, and then click the plus sign.

Reading is easier, too, in the new reading view. You can collapse parts of the document and focus on the text you want. If you need to stop reading before you reach the end, Word remembers where you left off, even on another device.



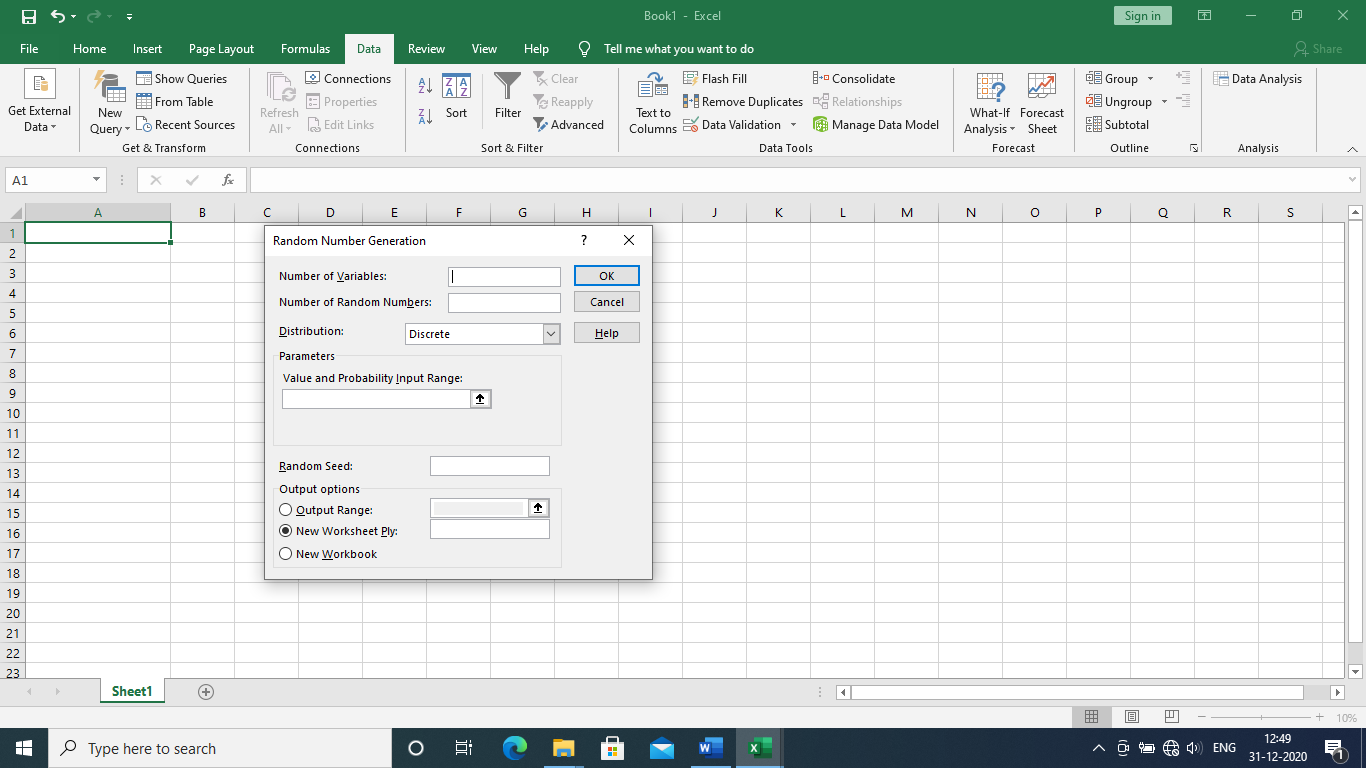
**Figure 18.11.** Screenshot for the Sequence of Clicks of Buttons, Namely Formulas → Math & Trig → RAND( )



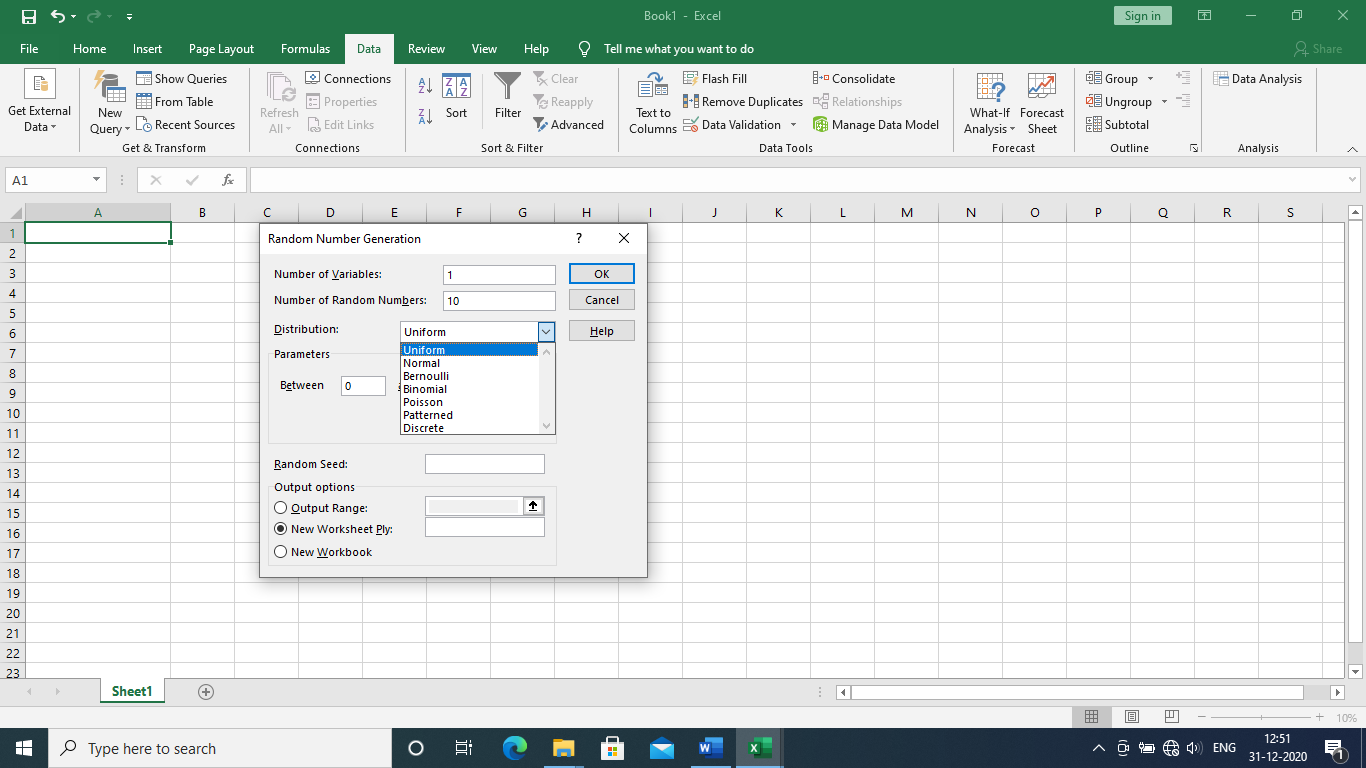
**Figure 18.12.** Screenshot after Clicking OK Button in the Drop-down Menu of Figure 18.11

## 18.3.2. Random Number Generation as Per Probability Distribution

The clicks of the sequence of buttons, namely *Data => Data Analysis => Random Number Generation* give the screenshot as shown in Figure 18.13. The click of sub-screen against distribution gives a screenshot as shown in Fig.18.14.



**Figure 18.13.** Screenshot for Clicks of Buttons, Namely Data => Data Analysis => Random Number Generation

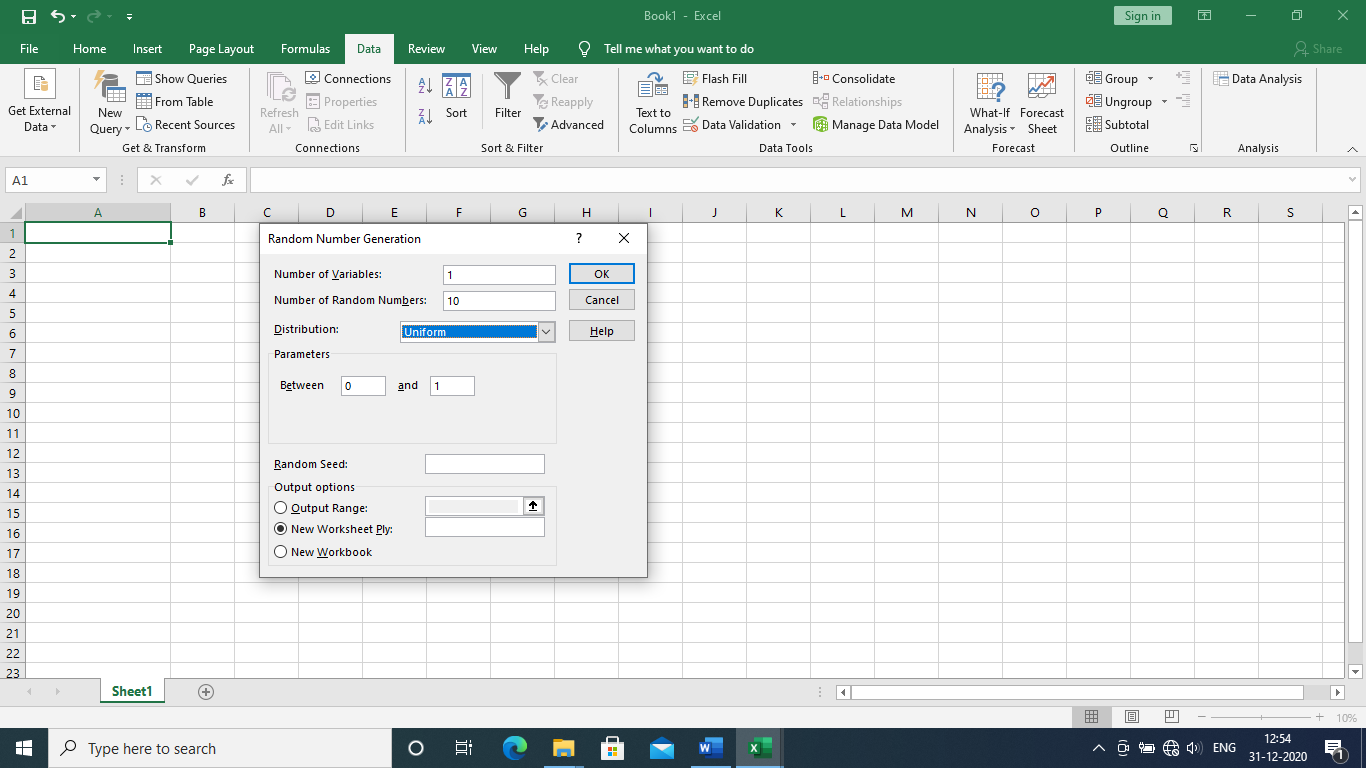


**Figure 18.14.** Screenshot after Clicking Sub-screen against Distribution in Figure 18.13

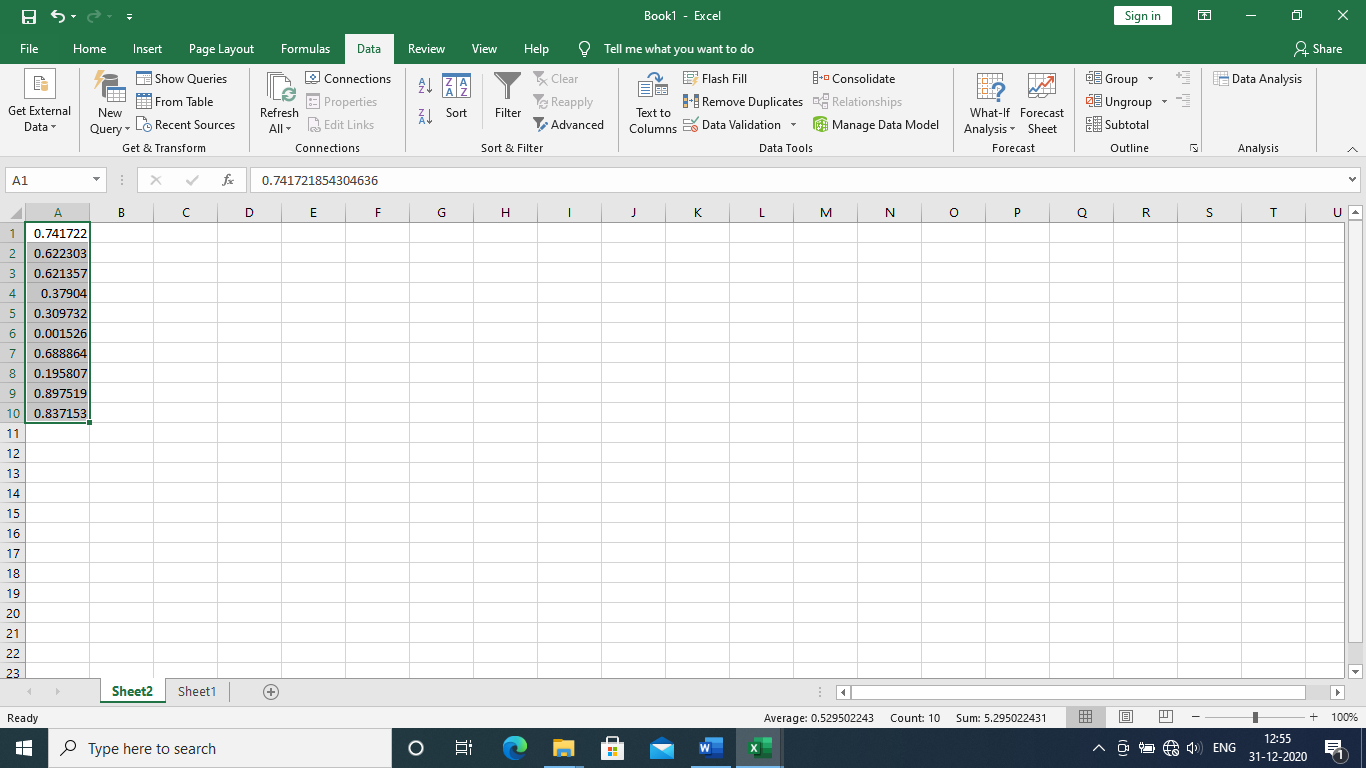
The selection of distribution under distribution and the corresponding drop-down menus are shown in different figures as summarized in Table 18.2.

**Table 18.2.** Summary of Figures for Selection Under Distribution

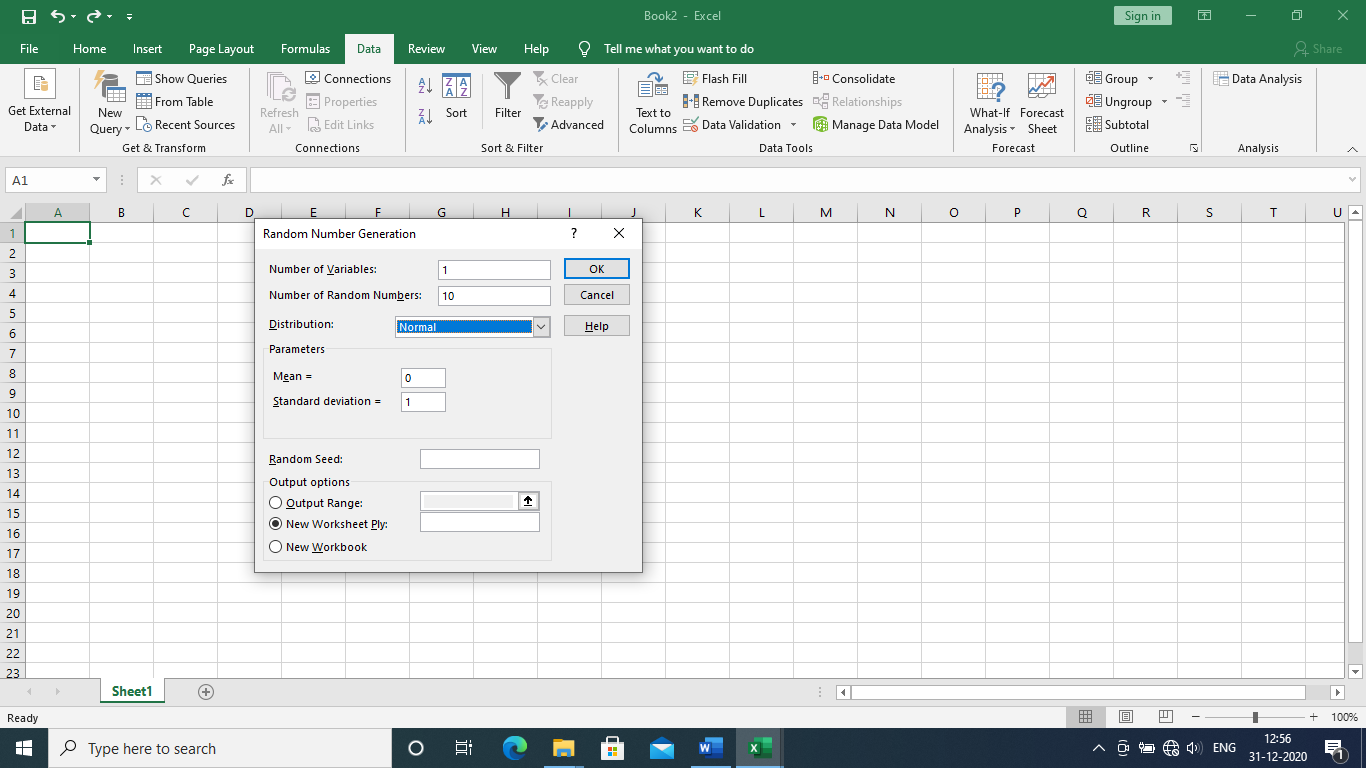
|  |  |  |  |
| --- | --- | --- | --- |
| S. No. | Name of Distribution | Name of Figure | Figure Number |
| 1 | Uniform | Screenshot of drop-down down menu for selecting Uniform | Figure 18.15 |
|  |  | Screenshot of random numbers | Figure 18.16 |
| 2 | Normal | Screenshot of drop-down down menu for selecting Normal | Figure 18.17 |
|  |  | Screenshot of random numbers | Figure 18.18 |
| 3 | Poisson | Screenshot of drop-down down menu for selecting Poisson | Figure 18.19 |
|  |  | Screenshot of random numbers | Figure 18.20 |
| 4 | Discrete | Screenshot of drop-down down menu for selecting Discrete | Figure 18.21 |
|  |  | Screenshot of random numbers | Figure 18.22 |



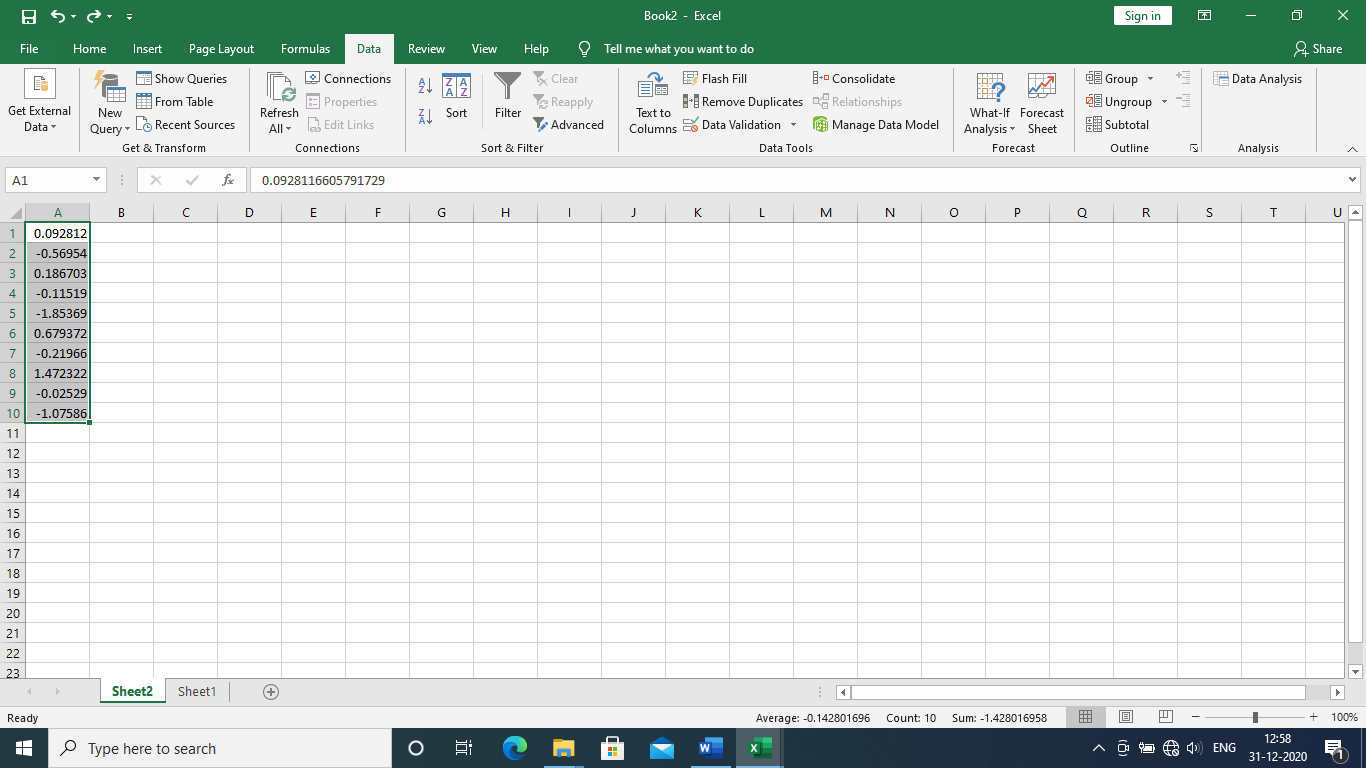
**Figure 18.15.** Screenshot after Selection of Uniform under Distribution



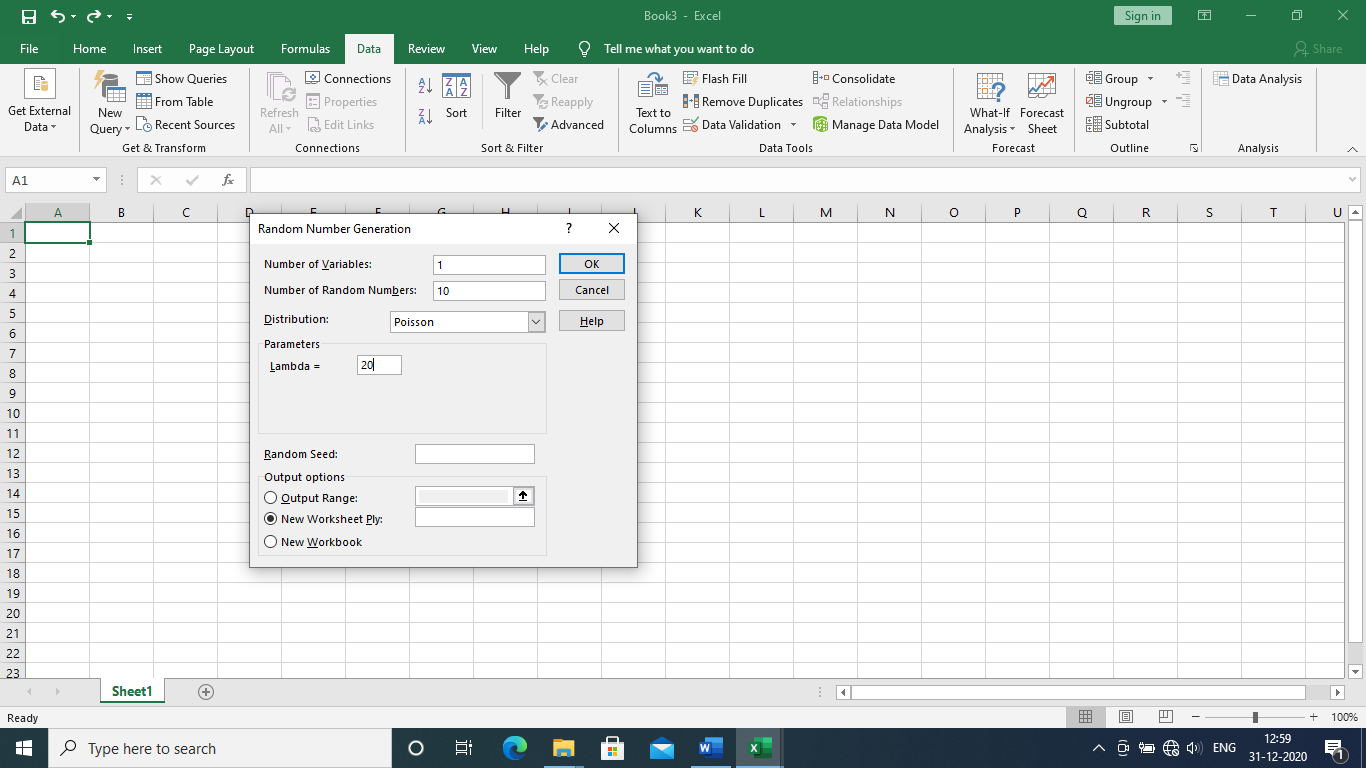
**Figure 18.16.** Screenshot of Uniformly Distributed Random Numbers



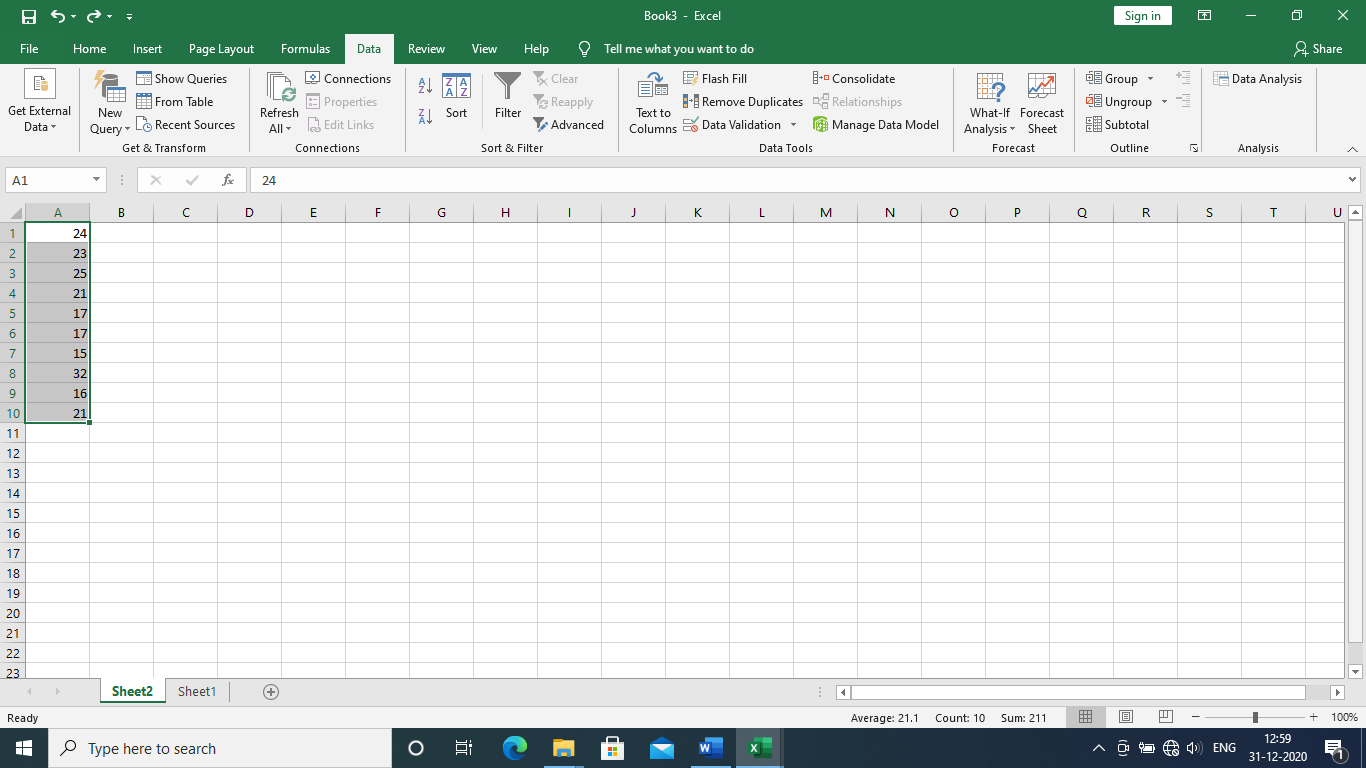
**Figure 18.17.** Screenshot after Selection of Normal under Distribution



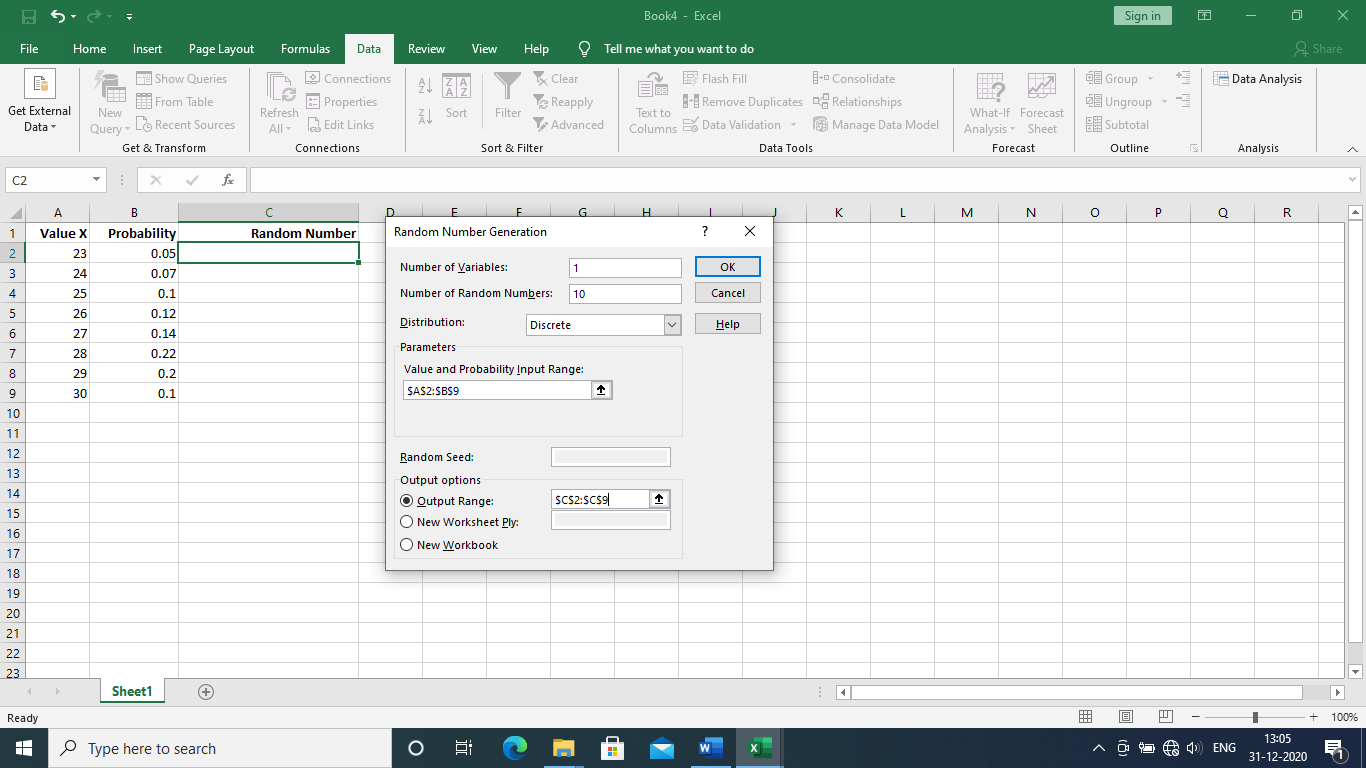
**Figure 18.18.** Screenshot of Normally Distributed Random Numbers



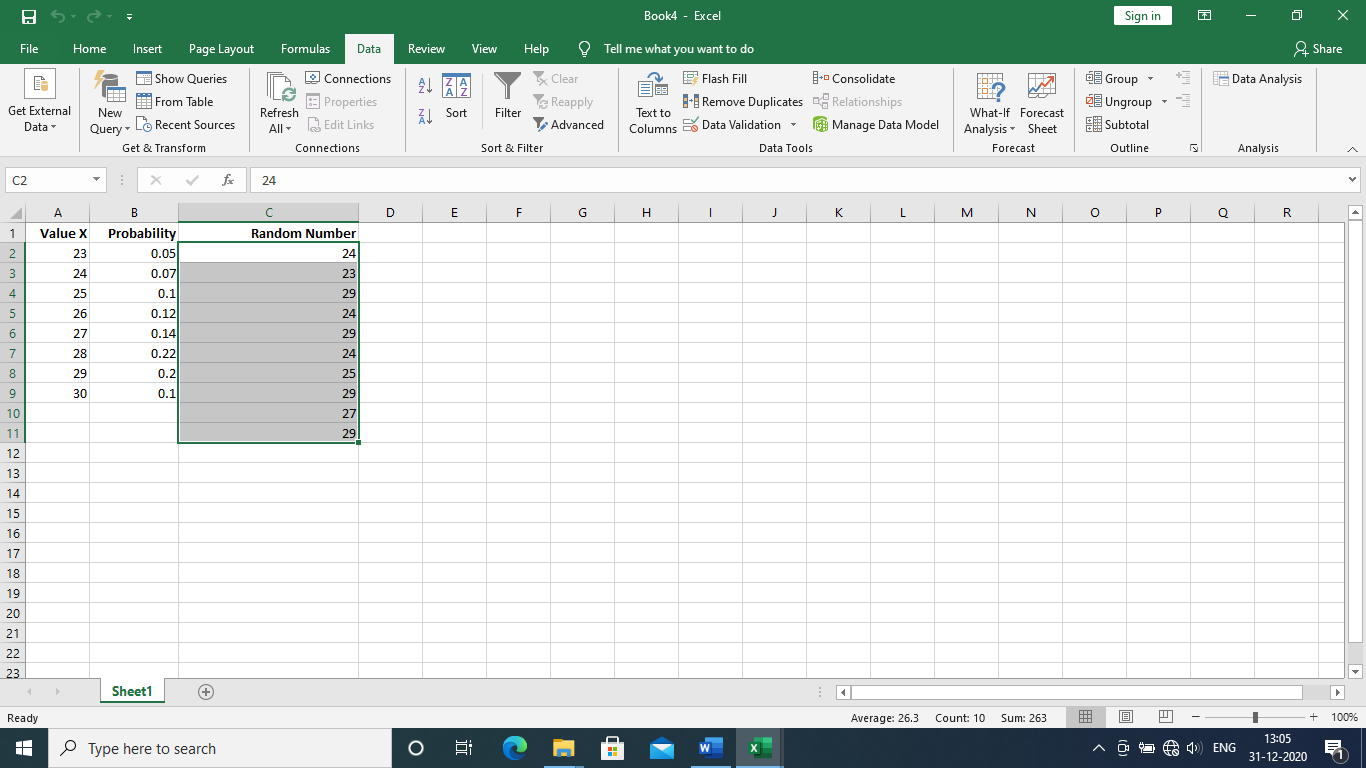
**Figure 18.19.** Screenshot after Selecting Poisson Distribution under Distribution and Entering Value for Lamda



**Figure 18.20.** Screenshot of Random Numbers as per Poisson Distribution



**Figure 18.21.** Screenshot after Selecting Discrete under Distribution and Filling the Data Items   
in the Drop-down Menu



**Figure 18.22.** Screenshot of Random Numbers as per Discrete Distribution

# 18.4. Monte-Carlo Simulation Using Excel Sheet

Monte-Carlo is a code word used in defence. Monte-Carlo simulation is a tool, which is used to estimate the occurrence of events in a probabilistic environment. Consider the estimation of the demand of a product for 10 months if the demand follows uniform distribution.

Let the probability distribution of the demand function of the product be as given below, which is a uniform distribution.

Where *a* sis the lower limit of the demand and *b* is the upper limit of the demand.

*X* is the random variable representing the demand.

The random variate to estimate *X* is given by the following formula.

If *a* and *b* are 400 and 600, respectively, then the uniform distribution of the demand of the product is as given below.

So the formula for the random variate, which estimates the demand of the product is as given below.

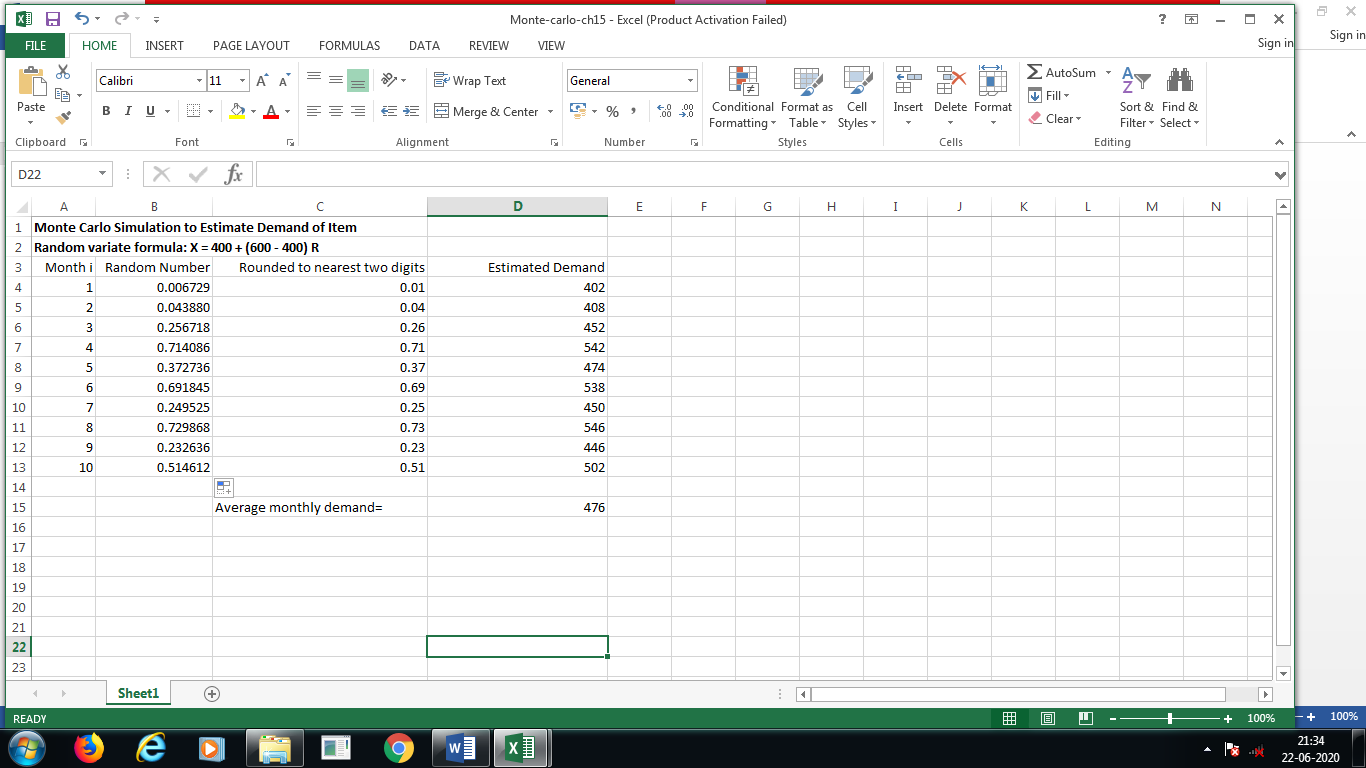
Where *X* is the random variate, which is the random variable, *R* is a uniformly distributed random number.

The simulation of the demand of the product for 10 months is shown in the screenshot of Figure 18.23. The formulas for the workings in Figure 18.23 are shown in Figure 18.24. The steps of simulating the demand values as shown in Figure 18.23 are listed below.

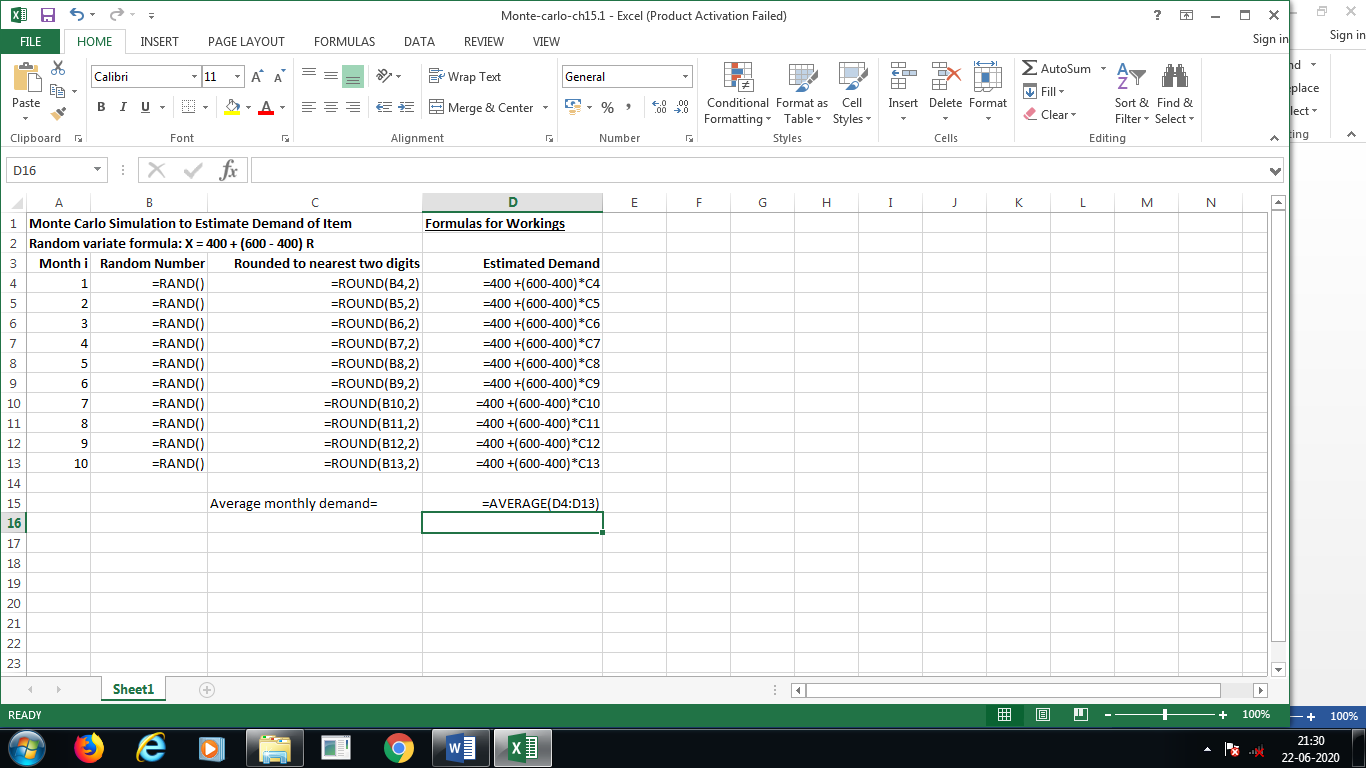
* Generate uniformly distributed random numbers using RAND( ) function as shown from cells B4 to B13.
* Round the random numbers in the cells from B4 to B13 to the nearest number with two decimal places as shown from C4 to C13 using ROUND function as shown below.
* Estimate the demand value for the Month i as shown from cells D4 to D13 using the following formula.

Demand of month i (*X*i) = 400 + (600 – 400) × Cell contain the random number for month i.

* Then find the average monthly demand using AVERAGE function as shown below. The average monthly demand is shown in cell D15.



**Figure 18.23.** Screenshot of the Results of Monte-Carlo Simulation to Estimate the Demand



**Figure 18.24.** Screenshot of Formulas of the Workings of Monte-Carlo Simulation to Estimate the Demand

**Summary**

* A real-life entity with all its members is called a population.
* A subset of a population for studying about the population is called a sample.
* In Excel, the method of selecting the samples may be either periodic or random. Normally, the random selection method is used.
* The clicks of buttons for sampling function are Data => Data Analysis => Sampling.
* If the sampling method is selected as periodic, then the data for "Period" is to be filled.
* Random sampling generates samples from a given population using simulation.
* Periodic sampling generates samples at a fixed interval from a given population.
* If the sampling method is selected as random, then "Number of Samples" is to be given as data for this option.
* The number of cells in the output range in the drop-down menu of the sampling function should be more than the number of samples to be generated, because there will be repetition of the samples in the generated samples.
* RAND function gives uniformly distributed random numbers.
* Monte-Carlo was a code word used in defence.
* Monte-Carlo simulation is used to estimate the values of the random variable of a probability distribution.
* The formula for the uniform distribution is

*P(X*) = 1/(b-a), where a ≤ *X* ≤b

Where,

*X* is the random variable

*a* is the lower limit of *X*

*b* is the upper limit of *X*

* The formula for the random variate of the uniform distribution is

*X* = *a* + (*b* – *a*) *R*

Where

*X* is the random variable

*R* is a uniformly distributed random number

*a* is the lower limit of *X*

*b* is the upper limit of *X*.

**Keywords**

**Monte-Carlo simulation** is used to estimate the values of the random variable of a probability distribution.

**Monte-Carlo** was a code word used in defence.

**Number of samples** in the sampling function available in Excel means the number of samples to be generated.

**Period** in the sampling function available in Excel means the number of periods used to generate samples.

**Periodic sampling** generates samples at a fixed interval from a given population.

**Population** is a real-life entity with all its members.

**Random sampling** generates samples from a given population using simulation.

**Sample** is a subset of that population for studying about the population.

**Exercise**

1. **Multiple-choice Questions**

Refer to SAGE website

1. **Review Questions**
2. Give the sequence of clicks from HOME for the SAMPLE function in Excel.
3. Explain the items to be filled in the drop-down menu of the SAMPLE function in Excel.
4. An investigator has to select 5 students from among 15 employees in a class. The register numbers and the names of the students are shown in the following table.

Select five students randomly from among 15 students in the class using the SAMPLE function in Excel.

|  |  |
| --- | --- |
| Employee Code | Name of Student |
| 101 | Arthi |
| 102 | Anbu |
| 103 | Beena |
| 104 | Domnic |
| 105 | Elango |
| 106 | Farook |
| 107 | Ganesh |
| 108 | Gopu |
| 109 | Hendry |
| 110 | Ilamparuthi |
| 111 | Johnson |
| 112 | Kumar |
| 113 | Lingam |
| 114 | Manimaran |
| 115 | Sundaram |

1. In an acceptance sampling plan, the quality assistant wants to select a sample with 20 units. The units of the population are numbers from 1 to 200. Select 20 units randomly to have 20 samples from among 200 units of the population using the sampling function of Excel.
2. There are 75 important Part B questions for a subject. In the last minute, a student of the subject requires 5 questions from among these 75 questions to revise. Select five PART B questions through random sampling using the sampling function of Excel.
3. Illustrate the process of generating 10 samples out of 40 observations through periodic sampling using the sampling function in Excel.
4. Illustrate the use of RAND function in Excel.
5. What is random variate of a probability distribution? Explain its use.
6. Give the formula for the random variate of uniform distribution and explain its components.
7. Refer to books and journals, and give the random variate formulas of the following probability distributions.
8. Exponential distribution
9. Poisson distribution
10. The demand of a product follows uniform distribution, whose formula is as follows.

*P(X* )= 1/(900 – 700) for 700 ≤ *X* ≤ 900

Simulate 20 demand values and find the average demand of the product using Monte-Carlo simulation in an Excel sheet.

1. Generate 30 random numbers as per normal distribution using Excel.
2. Generate 20 random numbers as per the following discrete distribution using Excel.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Demand | 200 | 210 | 220 | 230 | 240 | 250 | 260 |
| Probability | 0.05 | 0.10 | 0.15 | 0.16 | 0.20 | 0.25 | 0.09 |

1. Generate 20 random numbers distributed as per Poisson distribution with the value of its parameter Lamda as 12.

**References**

Adam. 2018. ‘How and Why of Running a Simulation in Excel’. *Spreadsheet Web*. Available at <https://www.spreadsheetweb.com/simulation-in-excel/> (accessed on 25 June 2020).

Check Your Assumptions. n.d. ‘How to Create a Random Sample in Excel (in 3 Minutes!)’. YouTube. Available at <https://www.youtube.com/watch?v=q8fU001P2lI> (accessed on 25 June 2020).

Dalgleish, Debra. n.d. ‘Excel RAND and RANDBETWEEN Functions’. *Contextures*. Available at <https://www.contextures.com/excelfunctionsrandrandbetween.html> (accessed on 8 July 2020).

DataScience Made Simple. n.d. *Data Sampling in Excel—Create Random Sample in Excel*. Available at <http://www.datasciencemadesimple.com/sampling-in-excel-create-random-sample-in-excel/> (accessed on 8 July 8, 2020).

Exceljet. n.d. ‘Excel RAND Function’. Available at <https://exceljet.net/excel-functions/excel-rand-function> (accessed on 8 July 2020).

Gordon, Geoffrey. 1969. *System Simulation*. Englewood Cliffs, NJ: Prentice Hall.

Naylor, T. J., J. L. Balintfy, D. S. Burdick, and K. Chu. 1966. *Computer Simulation Techniques*. New York, NY: John Wiley & Sons.

Nelson, S. L., and E. C. Nelson. 2016. ‘How to Sample Data in Excel’. *Dummies*. Available at <https://www.dummies.com/software/microsoft-office/excel/how-to-sample-data-in-excel/> (accessed on 24 June 2020).Panneerselvam, R. 1980. *Simulation Model for a Season Ticket Issue Counter* (Unpublished M.E. [Industrial Engineering] Mini-thesis). Chennai: Anna University, 1980.

Panneerselvam, R., and M. Veerapandian, M. 1990. ‘Simulation of a Season Ticket Issue Counter of a Transport Corporation’. *Industrial Engineering Journal* 19 (7): 21–25.

Panneerselvam, R., and P. Senthilkumar. 2013. *System Simulation Modelling and Languages*. New Delhi: PHI Learning Private Limited.

Scriber, T. J. 1974. *Simulation Using GPSS*. New York, NY: John Wiley & Sons.

Zaiontz, Charles. n.d. ‘Sampling’. *Real Statistics Using Excel*. Available at <http://www.real-statistics.com/sampling-distributions/sampling/> (accessed on 8 July 2020).