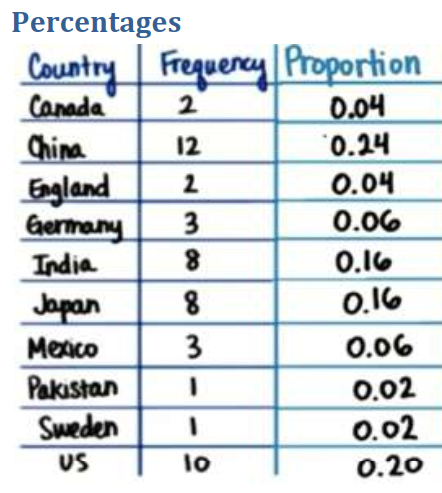
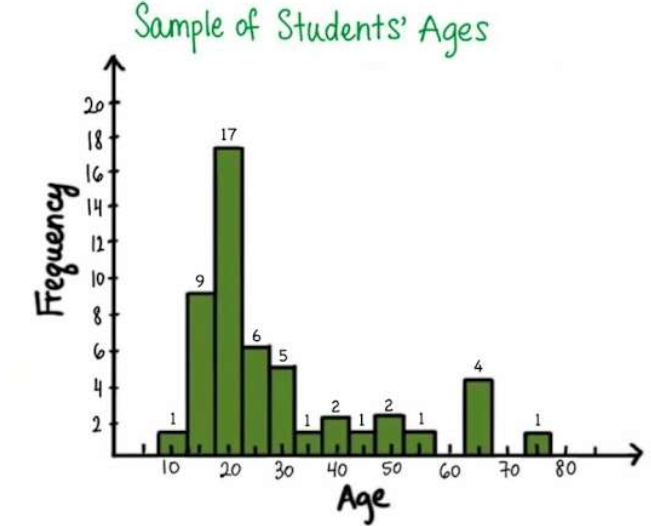
**Solutions Statistic Essentials**

**1**.



**2**. Please create manually.



**3. What is the probability of randomly selecting a person having age between 10 years and 20 years?**

27/50

**4. What is the probability of randomly selecting a person in the age 5yrs to 80 years?**

100%

**5. Create the prob density curve for the histogram**

<hint: joint the midpoints of each bar>

We’ll be discussing more about prob density curve in visualization class

6. **Find the mean, median and mode age of students**

Mode Age = 17.5 – 22.5

Mean Age = (10 X 1 + 9 X 15 + 17 X 20 + 6 X 25 + 5 X 30 + 1 X 35 + 2 X 40 + 1 X 45 + 2 X 50 + 1 X 55 + 4 X 65 + 1 X 75)/50 = 28.7 Approx (approx because we are approximating the age value to bucket mean, which is just an assumption. Eg. We have taken 10 yrs as the age of the only person in bucket 7.5 yrs to 12.5 yrs. This is just an assumption. )

Median Age = the age of the middle person i.e (t25 + t26)/2. Both t25 and t26 should lie in bucket 17.5 - 22.5 yrs.

**7. Is this distribution normal in shape? If not find out if it is right skewed or left skewed**

Not Normal, Right Skewed

**8**. **Find the range and IQR for above distribution.**

Range = 77.5 - 7.5 = 70

IQR = Q3 – Q1

Q3 is where 75% of values are covered i.e 37.5 values are covered = approx. 30

Q1 is where 25% of values are covered i.e 12.5 values are covered = approx 20

IQR = approx (30 – 20) = approx 10

**9**. **Outliars are values greater than**

Q3 + 1.5 \* IQR = 30 + 1.5\*10 = 45

So we have approx 8 outliars

**10**. **Difference between the bar chart and histogram**

- Bar Chart is for categorical variables, while histogram is for continuous numerical variables.

- Shape of bar-chart is not important as we can change the order of the categories in a bar-chart. Where-as in-case of a histogram the shape is important and communicates about the properties of distribution.