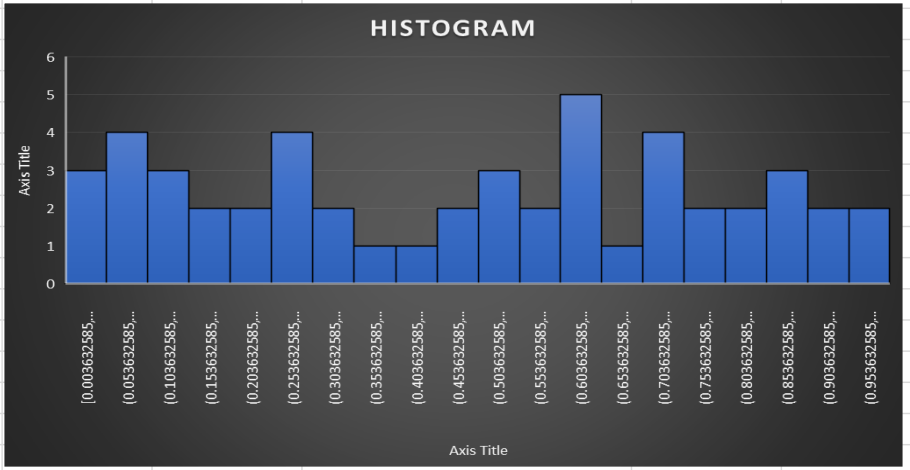
QUESTION 1		**Simulate 30 rolls with =RANDBETWEEN(1,6). What is the probability of rolling a 3 exactly 5 times? (Hint: Use BINOM.DIST)**							
ANSWER		Identify the binom.dist							
		Total rolls =30		Random bettween(1,6)					
		Numbers of trial n =30							
		probability of an even happening = **Numbers of ways it can happened/ Toatal numbers of outcomes**							
		probability of success rolling 3 on a one roll			**p= 1/6**				
		Numbers of sucess = X=5							

QUESTION 2 Generate 100 values in Excel using the continuous uniform distribution RAND() and plot a histogram. Describe the shape of the distribution

ANSWER Create random value using ran function

1	0.5058209015	51	0.2624980818
2	0.5032640703	52	0.4543559395
3	0.2426275784	53	0.2961709687
4	0.5266317788	54	0.3586398309
5	0.8115803033	55	0.5623002489
6	0.6523003376	56	0.4169649516
7	0.1760821561	57	0.238408503
8	0.4605324859	58	0.2314974791
9	0.7548541704	59	0.6137413885
10	0.1081774088	60	0.8768418443
11	0.4453867058	61	0.399515109
12	0.8244802699	62	0.4260689572
13	0.7510003423	63	0.8574980771
14	0.9953225413	64	0.3067116434
15	0.2954723437	65	0.5106078302
16	0.07813392372	66	0.195783871
17	0.7953361802	67	0.5701335525
18	0.6844916225	68	0.6905781831
19	0.1384111202	69	0.5307417254
20	0.7346731088	70	0.5908003708
21	0.9364984664	71	0.1983459286
22	0.7345952554	72	0.7320816038
23	0.4572478021	73	0.8571025035
24	0.510230344	74	0.05875990349
25	0.4498356875	75	0.308241415
26	0.4204289005	76	0.1058674849
27	0.1443153243	77	0.161287301
28	0.3719253369	78	0.5704417919
29	0.786424458	79	0.3917295491
30	0.2990964194	80	0.9376671082
31	0.9575359008	81	0.7407420621
32	0.01785749242	82	0.6381452307
33	0.6906851217	83	0.7926432111
34	0.0931117149	84	0.3521086277
35	0.9237178884	85	0.3730572709



		36	0.1007326894	86	0.4510732273						
		37	0.1032042425	87	0.4419456468						
		38	0.1496082863	88	0.2368493394						
		39	0.06961622391	89	0.6657059476						
		40	0.9413705941	90	0.364686759						
		41	0.5952541739	91	0.9243937052						
		42	0.8727820945	92	0.3374390785						
		43	0.8195319541	93	0.804941086						
		44	0.5691592015	94	0.1499439962						
		45	0.1176333908	95	0.6382843332						
		46	0.8206380786	96	0.4193960425						
		47	0.7624845604	97	0.4249524624						
		48	0.6735935766	98	0.7335337721						
		49	0.7888333482	99	0.455419553						
		50	0.6073064763	100	0.5122954557						

QUESTION3	A dataset has a mean of 50 and a standard deviation of 5. What percentage of values lie between 45 and 55 if the data follows a normal distribution?										
ANSWER	**MEAN=50**	**STD=5**									
	The interval 45 to 55										
	50+5=55										
	50-5=45		**68% values**								
	Approximately 68% of the data values lie between 45 and 55 when the data follows a normal distribution										

QUESTION 4	What is the concept of standardization (z-score), and why is it important in data analysis? Explain the formula and how standardization transforms a dataset.
ANSWER	Standardization is the process of converting raw data values into z-scores, which show how far and in what direction a value is from the mean, measured in standard deviation unit
	whether a value is above or below the mean
	how unusual or typical the value is compared to the rest of the data
	Z-score Formula
	standardized value (z-score)=original data value-mean of the dataset/standard deviation of the dataset
	When a dataset is standardized
	The new mean becomes 0
	The new standard deviation becomes 1
	The shape of the distribution stays the same
	Only the scale changes, not the relative positions

QUESTION5	What is Kurtosis and their type?	
ANSWER		Kurtosis is a statistical measure that describes the shape of a distribution, specifically the peakedness and thickness of the tails compared to a normal distribution.
		How sharp or flat the peak is
		How heavy or light the tails are (presence of extreme values)
	There are three main types of kurtosis	
	1)Mesokurtic	
	2)Leptokurtic	
	3) Platykurtic	
	1)Mesokurtic	Same kurtosis as a normal distribution
		Kurtosis value: 3
		Excess kurtosis: 0
		Moderate peak, normal tails
	2)Leptokurtic	Higher kurtosis than normal
		Kurtosis value: > 3
		Excess kurtosis: > 0
		Sharp peak and heavy tails
	3) Platykurtic	Lower kurtosis than normal
		Kurtosis value: < 3
		Excess kurtosis: < 0
		Flat peak and light tails

QUESTION 6	Explain why the uniform distribution is a good model for the outcome of rolling a fair die.
ANSWER	The uniform distribution is a good model for the outcome of rolling a fair die because each possible outcome has the same probability of occurring.
	Equal Probability of Each Outcome
	A fair die has six faces, numbered 1 through 6.
	P(1)=P(2)=P(3)=P(4)=P(5)=P(6)=6 1/6
	No number has a higher or lower chance of appearing.
	This equality of probabilities is the core requirement of a uniform distribution.
	Discrete Uniform Distribution
	Rolling a die is a discrete random experiment because
	The outcomes are countable
	Only whole numbers (1 to 6) are possible
	Fixed and Finite Sample Space
	The sample space of a die roll is:
	S={1,2,3,4,5,6}
	The set of outcomes is finite
	The outcomes do not change between trials
	Graphical Representation
	If we plot a histogram of a large number of die rolls:
	Each bar (1 to 6) has roughly the same height
	The distribution appears flat

QUESTION 7	Use Excel to compute the probability of getting at least 8 successes in 15 trials with success probability 0.5						
ANSWER	Number of trials: n=15n						
	Probability of success: p=0.5p						
	We want: P(X≥8)						
	P(X>8)=0.500						
	p=0.5, the distribution is symmetric, so the probability is very close to 0.5.)						

QUESTION 8	How does log transformation help in stabilizing variance and making data more normally distributed
ANSWER	A log transformation is a mathematical transformation applied to data to address two common problems in
	real-world datasets: non-constant variance (heteroscedasticity) and non-normality (skewness). Below is a clear, detailed explanation.
	What is a Log Transformation
	A log transformation replaces each data value xxx with its logarithm:
	y=log(x)
	log10(x) (base-10 log)
	ln(x) (natural log)
	How Log Transformation Stabilizes Variance
	Heteroscedasticity
	Small values have low variance
	Large values have high variance
	Spread increases as the mean increases
	Logarithms compress large values more than small values
	Reduces the influence of extreme values