

UNIVERSITY OF PETROLEUM & ENERGY STUDIES, DEHRADUN

Program	MCA	Semester	I
Course	Introduction to Data Science	Course Code	CSDS 7001

SN	Questions	COs	Marks																				
1	<p>The following table gives the frequency distribution of the weekly wages (in '00 Rs.) of 100 workers in a factory.</p> <table><tr><td>Weekly Wages ('00 Rs.)</td><td>20-24</td><td>25-29</td><td>30-34</td><td>35-39</td><td>40-44</td><td>45-49</td><td>50-54</td><td>55-59</td><td>60-64</td></tr><tr><td>Number of Workers</td><td>14</td><td>15</td><td>22</td><td>33</td><td>41</td><td>20</td><td>18</td><td>15</td><td>12</td></tr></table> <p>Draw the histogram and frequency polygon of the distribution.</p>	Weekly Wages ('00 Rs.)	20-24	25-29	30-34	35-39	40-44	45-49	50-54	55-59	60-64	Number of Workers	14	15	22	33	41	20	18	15	12	CO3	10
Weekly Wages ('00 Rs.)	20-24	25-29	30-34	35-39	40-44	45-49	50-54	55-59	60-64														
Number of Workers	14	15	22	33	41	20	18	15	12														
2	<p>Define the feature engineering. Explain the following terms:</p> <p>(i) Importance of feature engineering.</p> <p>(ii) Techniques used for feature engineering for Machine Learning,</p> <p>(iii) Best tools for feature engineering.</p>	CO4	10																				
3	<p>Write a short note on Decision Tree Algorithm. List down the attribute selection measures used by the ID3 algorithm to construct a decision tree.</p>	CO4	10																				
4	<p>Write a short note on cluster analysis.</p>	CO5	10																				
5	<p>Define the partition methods. Explain the K- means clustering algorithm with an example.</p>	CO5	10																				