**JAIPURIA INSTITUTE OF MANAGEMENT, INDORE**

**PGDM**

**FOURTH TRIMESTER (Batch 2021-23)**

**END TERM IMPROVEMENT EXAMINATION, DEC-2022**

|  |  |  |  |
| --- | --- | --- | --- |
| Course Name | Python for Business Analytics | Course Code | 40823 |
| Max. Time | **2 hours** | Max. Marks | **40** |

**INSTRUCTIONS:**

* *Please write the analysis in the jupyter notebook using multiline comments.*
* *Submit on one jupyter notebook file, with proper heading of each question*
* *For Q3(a and b write the interferences in the jupyter notebook mentioning the question number.*

**Questions.1** *The dataset contains data about various Bollywood movies (Bollywood.xls)* **(15 Marks)**

Release timings: HS (Holiday Season), LW (Long Weekend), N (Normal Days), FS (Festival Season).

List down the records present in the dataset? Explain the metadata information of the given dataset.

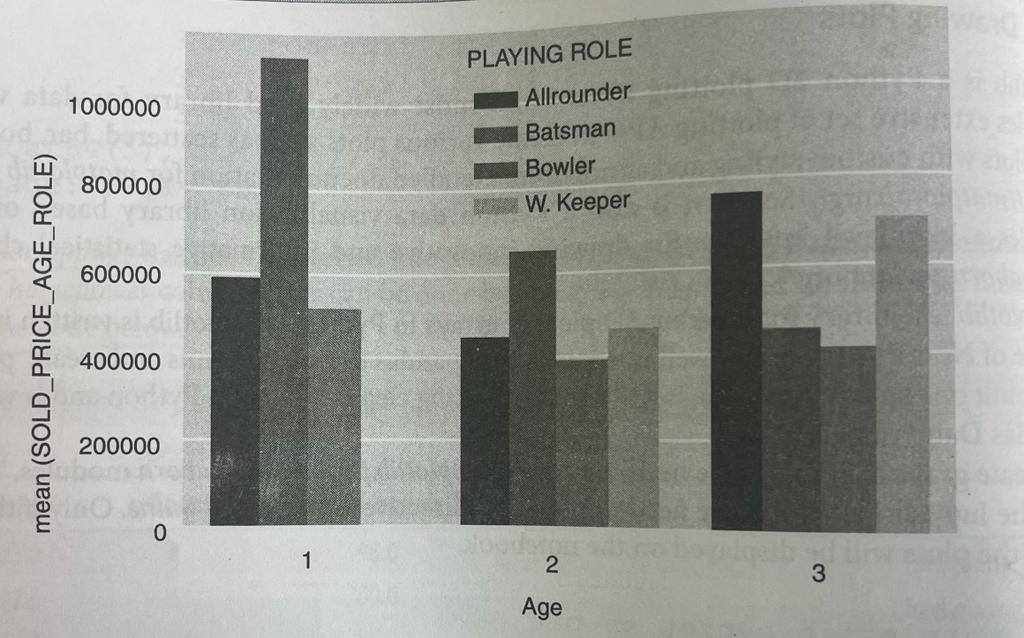
1. Explain with the help of suitable plot the number of the persons movies released during different release timings, is movies release more during LW or HS as compare to N days. Comment.
2. Comment on correlation of YouTubeviews with boxofficecollection? Choose appropriate plot to show the relation.
3. Compare the distribution of boxofficecollection as per different Genre. Can you interpret the results?
4. Derive the new column called Earning from where movies falling in different segments are categorized as below, as big, moderate and small. (Take saving amount as per your understanding).
5. Compare distribution of *budget* for different movies. Comment

**Questions.2** a) IIT preparation coaching center would like to understand the success of their students in term of clearing the IIT examination. The faculty members have been stating that the percentage of success is not more than 70%. They would like to test whether the claim made by their faculty members is correct. They picked up a random sample of 20 students and found 18 got selected. Write a code to test the hypothesis that the percentage of selection is less than 70%.

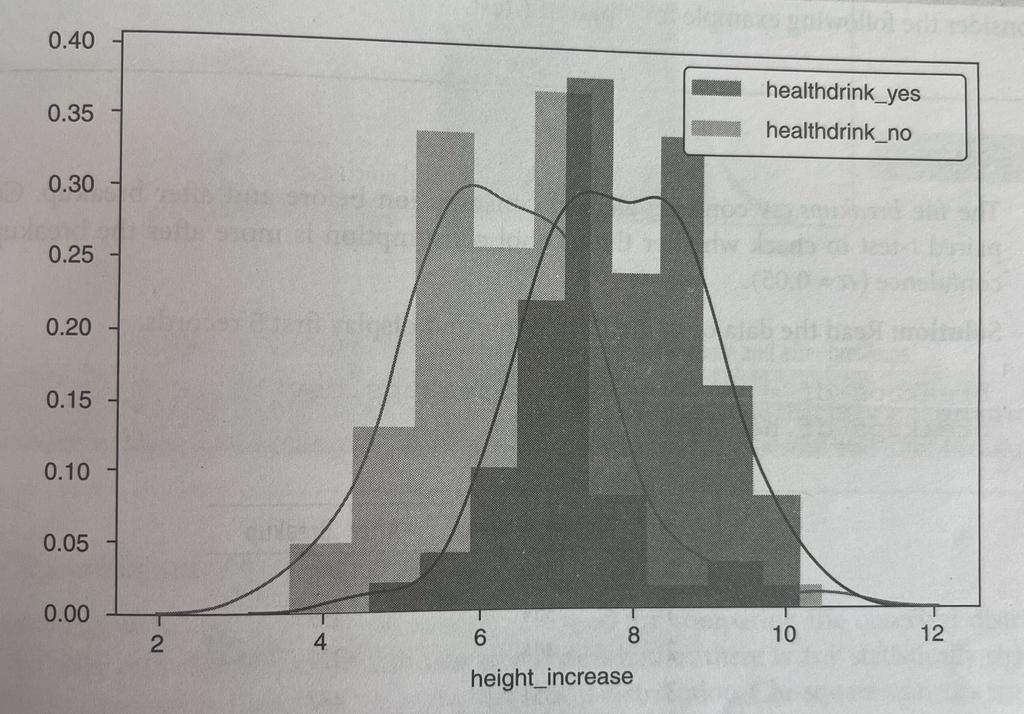
Sample of 20 [ 16.0, 16.0, 30.0, 37.0, 25.0, 19.0, 35.0, 27.0, 32.0, 34.0, 28.0, 24.0, 35.0, 24.0, 21.0, 32.0, 29.0, 24.0, 35.0, 28.0] **(5 Marks)**

1. In order to attract more students aspiring for admission into various international universities in terms of admission to ILTES coaching, Communication Gig Classes wants to design a strategy. Please assist them in developing the code that will allow the marketing team to devise a discount strategy based on the length of coaching attended by any student. If the coaching is for a full year, they will get a 35% discount. If it is for 6 months, they are eligible for a 25% discount. If they join for three months, they will get a 10% discount. Anyone who joins the weekend or biweekly program is only eligible for a free T-shirt and pen. **(5 Marks)**

**Questions.3** Analyze your result using an appropriate plot in order to justify your result. **(15 Marks)**

****

1. Explain the above chart, in the Above figure, (Age 1- 20-30 years, 2- 35-45 years and 3- above 45 years), is there is any correlation between age of the payer and playing role.

****

1. Explain the above chart, is there is any correlation of increase in height after taking health drink or not.