|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Exams are approved when answering 85% of the questions correctly, conditional on the approval of the evalutor. | | | | | | | | | | | | | | | | | | | | | | |  |
| # | Skill code | Question type (for available q-types refer to this [link](https://help.sabacloud.com/sabacloud/help-learning/topics/help-assessment-question-types-supported-by-tests.html)) | Question | Question choices (if applicable) | Correct answer | Link to image | Comments from OLC | Question – Iteration 1 | Question choices (if applicable) – Iteration 1 | Correct answer – Iteration 1 | Link to dataset (if applicable) | Question – Iteration 2 | Question choices (if applicable) – Iteration 2 | Correct answer – Iteration 2 | Link to dataset (if applicable) | Question – Iteration 3 | Question choices (if applicable) – Iteration 3 | Correct answer – Iteration 3 | Link to dataset (if applicable) | Question – Iteration 4 | Question choices (if applicable) – Iteration 4 | Correct answer – Iteration 4 | Link to dataset (if applicable) |
| Part 1 – Questions  Duration 10 minutes  Cannot navigate backwards | | | | | | | | | | | | | | | | | | | | | | | |
| 8 | DS1-1 | Multiple choice | **Q1** Which variable from below most likely comes from a normal distribution?  [Link to png image](https://worldbankgroup.sharepoint.com/:i:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/images_olc/n8_0.png?csf=1&web=1&e=qpkmtI) | A) x  B) y  C) z | A |  |  | Which variable from below most likely comes from a uniform distribution?  [Link to png image](https://worldbankgroup.sharepoint.com/:i:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/images_olc/n8_1.png?csf=1&web=1&e=zSBb97) | A) x  B) y  C) z | C |  | Which variable from below most likely comes from a lognormal distribution?  [Link to png image](https://worldbankgroup.sharepoint.com/:i:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/images_olc/n8_2.png?csf=1&web=1&e=sQ5Eb5) | A) x  B) y  C) z | B |  | Which variable from below most likely comes from a left-skewed distribution?  [Link to png image](https://worldbankgroup.sharepoint.com/:i:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/images_olc/n8_3.png?csf=1&web=1&e=dswXgp) | A) x  B) y  C) z | B |  | Which variable from below most likely comes from a bimodal distribution?  [Link to png image](https://worldbankgroup.sharepoint.com/:i:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/images_olc/n8_4.png?csf=1&web=1&e=NMKzEO) | A) x  B) y  C) z | A |  |
| 16 | DS1-4 | Multiple choice | **Q2** What is the output of the following Python code?  x = [6, 28, 5, 26, 49]  print(x[1:-1]) | A) [6, 28, 5, 26, 49]  B) [28, 5, 26, 49]  C) [28, 5, 26]  D) [6, 28, 5, 26] | C |  |  | What is the output of the following Python code?  x = [8, 9, 6, 10, 12, 14]  print(x[2:4]) | A) [9, 6, 10]  B) [6, 10]  C) [6, 10, 12]  D) [9, 6] | B |  | What is the output of the following Python code?  x = [8, 9, 12, 14]  print(x[3:]) | A)[14]  B) [12, 14]  C) [] | A |  | What is the output of the following Python code?  x = [4, 7, 29, 30, 18, 29]  print(x[3:-1]) | A) [30,18]  B) [29, 30, 18]  C) [29, 30, 18, 29]  D) [30, 18, 29] | A |  | What is the output of the following Python code?  x = [4, 7, 2, 3, 60, 18, 29]  print(x[:-1]) | A) [4, 7, 2, 3, 60, 18]  B) [4, 7, 2, 3, 60, 18, 29]  C) [4, 7, 2, 3, 60] | A |  |
| 18 | DS1-4 | All that apply | **Q3** Suppose you have defined the following list in your Python environment:  x = ['a', 'b', 'c']  Which of the following syntaxes would make x equal to:  ['a', 'b', 'c', 'd'] | A) x.append('d')  B) x += ['d']  C) x = x.append('d')  D) x = ['d'] | A, B |  |  | Suppose you have defined the following list in your Python environment:  x = [2, 3, 4]  Which of the following syntaxes would make x equal to:  [2, 3, 4, 5] | A) x.append(5)  B) x += 5  C) x = x.append('d')  D) x += [5] | A, D |  | Suppose you have defined the following list in your Python environment:  x = [2, 3, 4]  Which of the following syntaxes would make x equal to:  [1, 2, 3, 4] | A) x.insert(0, 1)  B) x.append(1, 0)  C) x = [1] + x  D) x = 1 + x | A, C |  | Suppose you have defined the following list in your Python environment:  x = [2, 3, 4]  Which of the following syntaxes would make x equal to:  [2, 3, 4, ‘a‘] | A) It is impossible, you cannot add a string to a list of integers  B) x.append(‘a’)  C) x = x + [‘a’]  D) x+= ‘a’ | B, C, D |  | Suppose you have defined the following list in your Python environment:  x = [1, 2, 3, 4]  Which of the following syntaxes would make x equal to:  [1, 2, 3] | A) x -= 1 B) x -= [4]  C) x = x[:-1]  D) x = x[0:-1] | C, D |  |
| 20 | DS1-4 | Multiple choice | **Q4** What is the output of the following Python code?   [Link to png image](https://worldbankgroup.sharepoint.com/:i:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/images_olc/n20_0.png?csf=1&web=1&e=p05cf3) | A) [True, True, False]  B) True  C) [False, False, True]  D) [False, False, True, True]  E) [False, False, True, False]  F) [True, True] | B |  |  | What is the output of the following Python code?  [Link to png image](https://worldbankgroup.sharepoint.com/:i:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/images_olc/n20_1.png?csf=1&web=1&e=4lrIle) | A)[‘a’, True, 1, 1.0]  B) a  C) [False, True, False, False | B |  | What is the output of the following Python code?  [Link to png image](https://worldbankgroup.sharepoint.com/:i:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/images_olc/n20_2.png?csf=1&web=1&e=u03Dgv) | A) [2,3,4]  B) [2, 3, 4, 1]  C) None  D) [3, 4, 5] | B |  | What is the output of the following Python code?  [Link to png image](https://worldbankgroup.sharepoint.com/:i:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/images_olc/n20_3.png?csf=1&web=1&e=eIKtZJ) | A) (2, True)  B) (1, True)  C) [False, (2, True), (3, True), False] | A |  | What is the output of the following Python code?  [Link to png image](https://worldbankgroup.sharepoint.com/:i:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/images_olc/n20_4.png?csf=1&web=1&e=NbL0un) | A) None  B) {1: 2, 2: 'a', 3: 'a', 4: 2}  C) {'a':2, 2:1, 3:1, 4:2} | B |  |
| 24 | DS1-5 | Multiple choice | **Q5** Suppose you have the below Pandas dataframe and code snippet. What are the columns that can be included in the resulting dataframe grouped\_df?    [Link to png image](https://worldbankgroup.sharepoint.com/:i:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/images_olc/n24_0.png?csf=1&web=1&e=4RrJOn) | A) [1, 2, 3, 4]  B) crop, Price, Quantity  C) Price, Quantity  D) farmerid, Price, Quantity  E) farmerid, crop, Price, Quantity | C |  |  | Suppose you have the below Pandas dataframe and code snippet. How many rows will grouped\_df have?  [Link to png image](https://worldbankgroup.sharepoint.com/:i:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/images_olc/n24_1.png?csf=1&web=1&e=qe8x3y) | A) 4  B) 5  C) 7  D) 8 | A |  | Suppose you have the below Pandas dataframe and code snippet. How many rows will grouped\_df have?  [Link to png image](https://worldbankgroup.sharepoint.com/:i:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/images_olc/n24_2.png?csf=1&web=1&e=aucgTG) | A) 4  B) 7  C) 3  D) 5 | B |  | Suppose you have the below Pandas dataframe and code snippet. How many index levels will grouped\_df have?  [Link to png image](https://worldbankgroup.sharepoint.com/:i:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/images_olc/n24_3.png?csf=1&web=1&e=2PEayu) | A) 1  B) 2  C) 4  D) 3 | B |  | Suppose you have the below Pandas dataframe and code snippet. What is the value for index Onion in grouped\_df?  [Link to png image](https://worldbankgroup.sharepoint.com/:i:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/images_olc/n24_4.png?csf=1&web=1&e=ojuaOa) | A)[260, 190]  B)[190. 260]  C) 225  D)None | C |  |
| 31 | DS1-6 | Fill in the blanks | **Q6** To upload your local commits to a remote repository:  git \_\_\_\_\_\_\_\_\_\_ <remote\_name><branch\_name> |  | push |  |  | To show the current condition of your repository, including modified files and staged changes:  git \_\_\_\_\_\_\_\_\_\_ |  | status |  | To initialize a new Git repository in your current directory:  git \_\_\_\_\_\_\_\_\_\_ |  | init |  | To fetch the latest changes from the specified branch on the remote repository and merge them into your current local branch:  git \_\_\_\_\_\_\_\_\_\_ <remote\_name><branch\_name> |  | pull |  | To Download a new repository from a remote server to your local machine:  git \_\_\_\_\_\_\_\_\_\_ <repo\_url> |  | clone |  |
| 42 | DS1-9 | True or False | **Q7** According to the World Bank’s Personal Data Privacy:  Personal data shall not be protected by any technical or organizational safeguards and freely processed without any restrictions, regardless of authorization |  | False |  |  | According to the World Bank’s Personal Data Privacy:  Personal data shall be transferred to third parties without regard of legitimacy and with no consideration for the protection of personal data |  | False |  | According to the World Bank’s Classification and Control Policy:  Strictly Confidential, Confidential and Official Use Only are the only information classification categories available |  | True |  | According to the World Bank’s Data Quality Assurance:  Verifiability, Transparency, and Relevance are Data Quality Principles |  | True |  | According to the World Bank’s Quality Assurance:  Data quality assurance is a shared responsibility of all staff involved in the production of Bank Development Data |  | True |  |
| 43 | DS1-9 | True or False | **Q8** Restricted information is assigned a security classification based on the level of harm posed by unauthorized disclosure |  | True |  |  | Restricted information is assigned a security classification based on restrictions imposed by the Information Provider or Originator |  | True |  | Restricted information is assigned a security classification based on the document file extension |  | False |  | Restricted information is assigned a security classification based on the extension of the document |  | False |  | Restricted information is assigned a security classification based on the date the information was originated |  | False |  |
| 46 | DS1-10 | True or False | **Q9** You were provided with AWS credentials. Since every time you want to access the data you need to provide them, it is a good idea to hard code them in the notebook you are using for your task |  | False |  |  | You were provided with AWS credentials. In order to keep them safe, it is a good idea to store them as environmental variables |  | True |  | It is not necessary to have disk encryption if your computer already has a password for logging in |  | False |  | If you lose your key to decrypt your disk, you lose all the information that was stored in it. |  | True |  | Your colleague has not received their AWS credentials yet. It is fine to share yours with her because you are working on the same project |  | False |  |
| Part 2 – Task Based Questions  Duration: 110 minutes  Cannot navigate backwards  This part will be approved conditional on the review of the material submitted on the final question | | | | | | | | | | | | | | | | | | | | | | | |
|  |  |  | **A Rubric for the candidate:**  You are provided with a dataset that stores information about movies like their genre, budget, rating, etc.  You are also provided with a second dataset that has the credits for the movies from the first dataset.  Load the datasets and answer the following questions: [Movies](https://worldbankgroup.sharepoint.com/:x:/t/DevelopmentDataPartnershipCommunity-WBGroup/EQhLHzrn_G5GhR7N5e5fDrgBWVYCf9kNyt263nLXFOK2UQ?e=mSqIKx), [Credits](https://worldbankgroup.sharepoint.com/:x:/t/DevelopmentDataPartnershipCommunity-WBGroup/EX_6M3ych2hBiX52y8z5cZUBhCk5FDZ-QFlbUTBsqc0_Hg)  At the end of the exam, you will be requested to load your notebook.script with the code written to obtain the solutions. The code will be reviewed by the evaluator who will determine whether you pass the exam or not. |  |  |  |  | **B Rubric for the candidate:**  You are provided with a dataset that stores information about movies like their genre, budget, rating, etc.  You are also provided with a second dataset that has the credits for the movies from the first dataset.  Load the datasets and answer the following questions: [Movies](https://worldbankgroup.sharepoint.com/:x:/t/DevelopmentDataPartnershipCommunity-WBGroup/EXcQRi-9j5NKo7wvoYJterABnYIESDFbNapsWJ4AGrtJVg?e=aRJOeV), [Credits](https://worldbankgroup.sharepoint.com/:x:/t/DevelopmentDataPartnershipCommunity-WBGroup/EV3qOjqTm7pDhGo99XfbOFIBzgVBhe747Mpf0XVKRG6c0A?e=Sse6UV)  At the end of the exam, you will be requested to load your notebook.script with the code written to obtain the solutions. The code will be reviewed by the evaluator who will determine whether you pass the exam or not. |  |  |  | **C Rubric for the candidate:**  You are provided with a dataset that stores information about movies like their genre, budget, rating, etc.  You are also provided with a second dataset that has the credits for the movies from the first dataset.  Load the datasets and answer the following questions: [Movies](https://worldbankgroup.sharepoint.com/:x:/t/DevelopmentDataPartnershipCommunity-WBGroup/EUebz1mId3lApz714ztHtuUB-COfFupkLJq6-NPqZqoH6g?e=zjv5ht), [Credits](https://worldbankgroup.sharepoint.com/:x:/t/DevelopmentDataPartnershipCommunity-WBGroup/EfiZttStBKNKrfVgRRfskE4Beav-NSLbf2Uz-_0UQSZsSA?e=iczM49)  At the end of the exam, you will be requested to load your notebook.script with the code written to obtain the solutions. The code will be reviewed by the evaluator who will determine whether you pass the exam or not. |  |  |  | **D Rubric for the candidate:**  You are provided with a dataset that stores information about movies like their genre, budget, rating, etc.  You are also provided with a second dataset that has the credits for the movies from the first dataset.  Load the datasets and answer the following questions: [Movies](https://worldbankgroup.sharepoint.com/:x:/t/DevelopmentDataPartnershipCommunity-WBGroup/Eataya_dSX1Hj4fVISDRQAQBrJblybs5b_JH4T8l-vxRZg?e=vlL2NJ), [Credits](https://worldbankgroup.sharepoint.com/:x:/t/DevelopmentDataPartnershipCommunity-WBGroup/EX_nIVzikt1Ah5zABm_N08ABCigfg1zXUQ4R8F6lcR-otQ?e=NFPSok)  At the end of the exam, you will be requested to load your notebook.script with the code written to obtain the solutions. The code will be reviewed by the evaluator who will determine whether you pass the exam or not. |  |  |  | **E Rubric for the candidate:**  You are provided with a dataset that stores information about movies like their genre, budget, rating, etc.  You are also provided with a second dataset that has the credits for the movies from the first dataset.  Load the datasets and answer the following questions: [Movies](https://worldbankgroup.sharepoint.com/:x:/t/DevelopmentDataPartnershipCommunity-WBGroup/EQocsY7HuA1Bkn4UHtb8-UABy8zUUmxaRlIdSsZrkEEJjg?e=hHvce7), [Credits](https://worldbankgroup.sharepoint.com/:x:/t/DevelopmentDataPartnershipCommunity-WBGroup/EZIdfxP4rO9IgGdTziyIVh4BxHLklzVXj6pmqIgl1tUghA?e=80cwLu)  At the end of the exam, you will be requested to load your notebook.script with the code written to obtain the solutions. The code will be reviewed by the evaluator who will determine whether you pass the exam or not. |  |  |  |
| 34 | DS1-7 | Fill in the blanks | **Q10** There are \_\_\_\_\_\_\_\_\_\_ columns in movie dataset and \_\_\_\_\_\_\_\_\_\_ columns in credits dataset. |  | 20, 4 | [M0](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/movies_iter_0.csv?d=w3a1f4b08fce7466e851ecde5ee5f0eb8&csf=1&web=1&e=6m2f5h)  [C0](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/credits_iter_0.csv?d=w7c33fa7f879c4168897e76cbccf97195&csf=1&web=1&e=fC1Oj5) |  | **Q10** There are \_\_\_\_\_\_\_\_\_\_ columns in movie dataset and \_\_\_\_\_\_\_\_\_\_ columns in credits dataset. |  | 21,3 | [M1](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/movies_iter_1.csv?d=w2f4610778fbd4a93a3bc2fa1826d7ab0&csf=1&web=1&e=l6IkWp)  [C1](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/credits_iter_1.csv?d=w3a3aea5d9b9343ba846a3df577db3852&csf=1&web=1&e=zj2B74) | **Q10**There are \_\_\_\_\_\_\_\_\_\_ columns in movie dataset and \_\_\_\_\_\_\_\_\_\_ columns in credits dataset. |  | 22,3 | [M2](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/movies_iter_2.csv?d=w59cf9b4777884079a73ef5e33b47b6e5&csf=1&web=1&e=pQ9SiD)  [C2](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/credits_iter_2.csv?d=wd4b699f804ad4aa3adf5604517ec904e&csf=1&web=1&e=tMGGNb) | **Q10** There are \_\_\_\_\_\_\_\_\_\_ columns in movie dataset and \_\_\_\_\_\_\_\_\_\_ columns in credits dataset. |  | 21,3 | [M3](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/movies_iter_3.csv?d=wafc95aab49dd477d8f87d52120d14004&csf=1&web=1&e=ehvJeq)  [C3](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/credits_iter_3.csv?d=w5c21e77f92e240dd879cc0066fcdd3c0&csf=1&web=1&e=O6ZKnv) | **Q10** There are \_\_\_\_\_\_\_\_\_\_ columns in movie dataset and \_\_\_\_\_\_\_\_\_\_ columns in credits dataset. |  | 22,3 | [M4](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/movies_iter_4.csv?d=w8eb11c0ab8c7410d927e141ed6fcf940&csf=1&web=1&e=7nR2Tc)  [C4](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/credits_iter_4.csv?d=w137f1d92acf848ef806753ce2c88561e&csf=1&web=1&e=zR40Fl) |
| 35 | DS1-7 | Fill in the blanks | **Q11** There are total \_\_\_\_\_\_\_\_\_\_ missing values in movie and \_\_\_\_\_\_\_\_\_\_ missing values in credit dataset |  | 6852, 0 | [M0](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/movies_iter_0.csv?d=w3a1f4b08fce7466e851ecde5ee5f0eb8&csf=1&web=1&e=6m2f5h)  [C0](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/credits_iter_0.csv?d=w7c33fa7f879c4168897e76cbccf97195&csf=1&web=1&e=fC1Oj5) |  | **Q11**There are total \_\_\_\_\_\_\_\_\_\_ missing values in movie and \_\_\_\_\_\_\_\_\_\_ missing values in credit dataset |  | 6502, 0 | [M1](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/movies_iter_1.csv?d=w2f4610778fbd4a93a3bc2fa1826d7ab0&csf=1&web=1&e=l6IkWp)  [C1](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/credits_iter_1.csv?d=w3a3aea5d9b9343ba846a3df577db3852&csf=1&web=1&e=zj2B74) | **Q11** There are total \_\_\_\_\_\_\_\_\_\_ missing values in movie and \_\_\_\_\_\_\_\_\_\_ missing values in credit dataset |  | 4830, 0 | [M2](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/movies_iter_2.csv?d=w59cf9b4777884079a73ef5e33b47b6e5&csf=1&web=1&e=pQ9SiD)  [C2](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/credits_iter_2.csv?d=wd4b699f804ad4aa3adf5604517ec904e&csf=1&web=1&e=tMGGNb) | **Q11** There are total \_\_\_\_\_\_\_\_\_\_ missing values in movie and \_\_\_\_\_\_\_\_\_\_ missing values in credit dataset |  | 4290, 0 | [M3](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/movies_iter_3.csv?d=wafc95aab49dd477d8f87d52120d14004&csf=1&web=1&e=ehvJeq)  [C3](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/credits_iter_3.csv?d=w5c21e77f92e240dd879cc0066fcdd3c0&csf=1&web=1&e=O6ZKnv) | **Q11** There are total \_\_\_\_\_\_\_\_\_\_ missing values in movie and \_\_\_\_\_\_\_\_\_\_ missing values in credit dataset |  | 4429, 0 | [M4](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/movies_iter_4.csv?d=w8eb11c0ab8c7410d927e141ed6fcf940&csf=1&web=1&e=7nR2Tc)  [C4](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/credits_iter_4.csv?d=w137f1d92acf848ef806753ce2c88561e&csf=1&web=1&e=zR40Fl) |
| 38 | DS1-7 | Fill in the blanks | **Q12** What would be the best way to impute missing values in the following columns :  1- vote\_average: \_\_\_\_\_\_\_\_\_\_  2- popularity: \_\_\_\_\_\_\_\_\_\_  Choose between mean, median, mode Note: Provide your answer using lowercase |  | 1- mean  2-median | [M0](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/movies_iter_0.csv?d=w3a1f4b08fce7466e851ecde5ee5f0eb8&csf=1&web=1&e=6m2f5h)  [C0](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/credits_iter_0.csv?d=w7c33fa7f879c4168897e76cbccf97195&csf=1&web=1&e=fC1Oj5) |  | **Q12**What would be the best way to impute missing values in the following columns :  1- runtime: \_\_\_\_\_\_\_\_\_\_  2- revenue: \_\_\_\_\_\_\_\_\_\_  Choose between mean, median, mode |  | 1- median  2-median | [M1](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/movies_iter_1.csv?d=w2f4610778fbd4a93a3bc2fa1826d7ab0&csf=1&web=1&e=l6IkWp)  [C1](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/credits_iter_1.csv?d=w3a3aea5d9b9343ba846a3df577db3852&csf=1&web=1&e=zj2B74) | **Q12** What would be the best way to impute missing values in the following columns :  1- runtime: \_\_\_\_\_\_\_\_\_\_  2- budget: \_\_\_\_\_\_\_\_\_\_  Choose between mean, median, mode |  | 1- median  2-median | [M2](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/movies_iter_2.csv?d=w59cf9b4777884079a73ef5e33b47b6e5&csf=1&web=1&e=pQ9SiD)  [C2](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/credits_iter_2.csv?d=wd4b699f804ad4aa3adf5604517ec904e&csf=1&web=1&e=tMGGNb) | **Q12** What would be the best way to impute missing values in the following columns :  1- vote\_average: \_\_\_\_\_\_\_\_\_\_  2- vote\_count: \_\_\_\_\_\_\_\_\_\_  Choose between mean, median, mode |  | 1- mean  2-median | [M3](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/movies_iter_3.csv?d=wafc95aab49dd477d8f87d52120d14004&csf=1&web=1&e=ehvJeq)  [C3](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/credits_iter_3.csv?d=w5c21e77f92e240dd879cc0066fcdd3c0&csf=1&web=1&e=O6ZKnv) | **Q12** What would be the best way to impute missing values in the following columns :  1- budget: \_\_\_\_\_\_\_\_\_\_  2- popularity: \_\_\_\_\_\_\_\_\_\_  Choose between mean, median, mode |  | 1- median  2-median | [M4](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/movies_iter_4.csv?d=w8eb11c0ab8c7410d927e141ed6fcf940&csf=1&web=1&e=7nR2Tc)  [C4](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/credits_iter_4.csv?d=w137f1d92acf848ef806753ce2c88561e&csf=1&web=1&e=zR40Fl) |
| 39 | DS1-7 | Fill in the blanks | **Q13** The analysis will continue by only using the records that do not have any missing values in the following columns: budget, popularity, revenue, runtime, vote\_average, vote\_count.  The new data frame has \_\_\_\_ rows |  | 2527 | [M0](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/movies_iter_0.csv?d=w3a1f4b08fce7466e851ecde5ee5f0eb8&csf=1&web=1&e=6m2f5h)  [C0](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/credits_iter_0.csv?d=w7c33fa7f879c4168897e76cbccf97195&csf=1&web=1&e=fC1Oj5) |  | **Q13**The analysis will continue by only using the records that do not have any missing values in the following columns: budget, popularity, revenue, runtime, vote\_average, vote\_count.  The new data frame has \_\_\_\_ rows |  | 2734 | [M1](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/movies_iter_1.csv?d=w2f4610778fbd4a93a3bc2fa1826d7ab0&csf=1&web=1&e=l6IkWp)  [C1](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/credits_iter_1.csv?d=w3a3aea5d9b9343ba846a3df577db3852&csf=1&web=1&e=zj2B74) | **Q13** The analysis will continue by only using the records that do not have any missing values in the following columns: budget, popularity, revenue, runtime, vote\_average, vote\_count.  The new data frame has \_\_\_\_ rows |  | 4022 | [M2](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/movies_iter_2.csv?d=w59cf9b4777884079a73ef5e33b47b6e5&csf=1&web=1&e=pQ9SiD)  [C2](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/credits_iter_2.csv?d=wd4b699f804ad4aa3adf5604517ec904e&csf=1&web=1&e=tMGGNb) | **Q13** The analysis will continue by only using the records that do not have any missing values in the following columns: budget, popularity, revenue, runtime, vote\_average, vote\_count.  The new data frame has \_\_\_\_ rows |  | 4489 | [M3](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/movies_iter_3.csv?d=wafc95aab49dd477d8f87d52120d14004&csf=1&web=1&e=ehvJeq)  [C3](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/credits_iter_3.csv?d=w5c21e77f92e240dd879cc0066fcdd3c0&csf=1&web=1&e=O6ZKnv) | **Q13** The analysis will continue by only using the records that do not have any missing values in the following columns: budget, popularity, revenue, runtime, vote\_average, vote\_count.  The new data frame has \_\_\_\_ rows |  | 4365 | [M4](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/movies_iter_4.csv?d=w8eb11c0ab8c7410d927e141ed6fcf940&csf=1&web=1&e=7nR2Tc)  [C4](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/credits_iter_4.csv?d=w137f1d92acf848ef806753ce2c88561e&csf=1&web=1&e=zR40Fl) |
| 3 | DS1-1 | Fill in the blanks | **Q14** What is the budget threshold or budget value (as an integer) below which 50% of the movie budgets are situated? |  | 14000000 | [M0](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/movies_iter_0.csv?d=w3a1f4b08fce7466e851ecde5ee5f0eb8&csf=1&web=1&e=6m2f5h)  [C0](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/credits_iter_0.csv?d=w7c33fa7f879c4168897e76cbccf97195&csf=1&web=1&e=fC1Oj5) |  | **Q14** What is the budget threshold or budget value (as an integer) below which 50% of the movie budgets are situated? |  | 40000000 | [M1](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/movies_iter_1.csv?d=w2f4610778fbd4a93a3bc2fa1826d7ab0&csf=1&web=1&e=l6IkWp)  [C1](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/credits_iter_1.csv?d=w3a3aea5d9b9343ba846a3df577db3852&csf=1&web=1&e=zj2B74) | **Q14** What is the `vote\_average` for the 10th percentile? |  | 4.9 | [M2](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/movies_iter_2.csv?d=w59cf9b4777884079a73ef5e33b47b6e5&csf=1&web=1&e=pQ9SiD)  [C2](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/credits_iter_2.csv?d=wd4b699f804ad4aa3adf5604517ec904e&csf=1&web=1&e=tMGGNb) | **Q14** What is the vote\_average that 90% of the movies have a value greater than it? |  | 7.3 | [M3](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/movies_iter_3.csv?d=wafc95aab49dd477d8f87d52120d14004&csf=1&web=1&e=ehvJeq)  [C3](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/credits_iter_3.csv?d=w5c21e77f92e240dd879cc0066fcdd3c0&csf=1&web=1&e=O6ZKnv) | **Q14** What is the vote\_average that 10% of the movies have a value lower than it? |  | 4.9 | [M4](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/movies_iter_4.csv?d=w8eb11c0ab8c7410d927e141ed6fcf940&csf=1&web=1&e=7nR2Tc)  [C4](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/credits_iter_4.csv?d=w137f1d92acf848ef806753ce2c88561e&csf=1&web=1&e=zR40Fl) |
| 13 | DS1-3 | Fill in the blanks | **Q15** Complete the following line of code to calculate the mean budget by original\_language:    movies.\_\_\_\_\_\_ ('original\_language')[“\_\_\_\_\_\_\_”].\_\_\_\_\_\_() Note: Provide your answer using lowercase |  | groupby  budget  mean | [M0](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/movies_iter_0.csv?d=w3a1f4b08fce7466e851ecde5ee5f0eb8&csf=1&web=1&e=6m2f5h)  [C0](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/credits_iter_0.csv?d=w7c33fa7f879c4168897e76cbccf97195&csf=1&web=1&e=fC1Oj5) |  | **Q15** Complete the following line of code to calculate the mean budget and mean vote\_average by original\_language:    movies.\_\_\_\_\_\_( “\_\_\_\_\_\_\_\_”)[[“budget”, “\_\_\_\_\_\_\_\_:]].mean()  Note: Provide your answer using lowercase |  | groupby  original\_language  vote\_average | [M1](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/movies_iter_1.csv?d=w2f4610778fbd4a93a3bc2fa1826d7ab0&csf=1&web=1&e=l6IkWp)  [C1](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/credits_iter_1.csv?d=w3a3aea5d9b9343ba846a3df577db3852&csf=1&web=1&e=zj2B74) | **Q15** Complete the following line of code to calculate the mean budget by genre:    Movies.\_\_\_\_\_\_\_('main\_genre')[“\_\_\_\_\_\_”].mean()  Note: Provide your answer using lowercase |  | groupby  budget | [M2](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/movies_iter_2.csv?d=w59cf9b4777884079a73ef5e33b47b6e5&csf=1&web=1&e=pQ9SiD)  [C2](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/credits_iter_2.csv?d=wd4b699f804ad4aa3adf5604517ec904e&csf=1&web=1&e=tMGGNb) | **Q15** Complete the following line of code to calculate the mean budget by genre:    movies.\_\_\_\_\_\_ ("\_\_\_\_\_\_\_\_")['budget'].mean()  Note: Provide your answer using lowercase |  | groupby  main\_genre | [M3](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/movies_iter_3.csv?d=wafc95aab49dd477d8f87d52120d14004&csf=1&web=1&e=ehvJeq)  [C3](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/credits_iter_3.csv?d=w5c21e77f92e240dd879cc0066fcdd3c0&csf=1&web=1&e=O6ZKnv) | **Q15** Complete the following line of code to calculate the mean budget by genre and original\_language:    movies.\_\_\_\_\_\_ ([“\_\_\_\_\_\_”, “\_\_\_\_\_\_”])['budget'].mean()  Note: Provide your answer using lowercase |  | groupby  main\_genre original\_language | [M4](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/movies_iter_4.csv?d=w8eb11c0ab8c7410d927e141ed6fcf940&csf=1&web=1&e=7nR2Tc)  [C4](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/credits_iter_4.csv?d=w137f1d92acf848ef806753ce2c88561e&csf=1&web=1&e=zR40Fl) |
| 14 | DS1-3 | Fill in the blanks | **Q16** Which original language had the second highest mean budget expenditure?  Note: Provide your answer using lowercase |  | ja | [M0](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/movies_iter_0.csv?d=w3a1f4b08fce7466e851ecde5ee5f0eb8&csf=1&web=1&e=6m2f5h)  [C0](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/credits_iter_0.csv?d=w7c33fa7f879c4168897e76cbccf97195&csf=1&web=1&e=fC1Oj5) |  | **Q16** Which original language had the lowest average vote\_average?  Note: Provide your answer using lowercase |  | ro | [M1](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/movies_iter_1.csv?d=w2f4610778fbd4a93a3bc2fa1826d7ab0&csf=1&web=1&e=l6IkWp)  [C1](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/credits_iter_1.csv?d=w3a3aea5d9b9343ba846a3df577db3852&csf=1&web=1&e=zj2B74) | **Q16** Which genre had the highest mean budget expenditure?  Note: Provide your answer using lowercase |  | Animation | [M2](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/movies_iter_2.csv?d=w59cf9b4777884079a73ef5e33b47b6e5&csf=1&web=1&e=pQ9SiD)  [C2](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/credits_iter_2.csv?d=wd4b699f804ad4aa3adf5604517ec904e&csf=1&web=1&e=tMGGNb) | **Q16** Which genre had the lowest mean budget expenditure?  Note: Provide your answer using lowercase |  | Foreign | [M3](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/movies_iter_3.csv?d=wafc95aab49dd477d8f87d52120d14004&csf=1&web=1&e=ehvJeq)  [C3](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/credits_iter_3.csv?d=w5c21e77f92e240dd879cc0066fcdd3c0&csf=1&web=1&e=O6ZKnv) | **Q16** Which genre and language had the second highest mean budget expenditure? Genre: \_\_\_\_\_ and language \_\_\_\_\_\_  Note: Provide your answer using lowercase |  | Fantasy, ko | [M4](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/movies_iter_4.csv?d=w8eb11c0ab8c7410d927e141ed6fcf940&csf=1&web=1&e=7nR2Tc)  [C4](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/credits_iter_4.csv?d=w137f1d92acf848ef806753ce2c88561e&csf=1&web=1&e=zR40Fl) |
| 15 | DS1-3 | Fill in the blanks | **Q17** The number of movies with the following characteristics is \_\_\_\_\_\_\_:  - Spent less than 500,000  - The runtime is above 120 minutes or below 100 minutes  -The original language is abbreviated as 'fr' or 'es' |  | 8 | [M0](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/movies_iter_0.csv?d=w3a1f4b08fce7466e851ecde5ee5f0eb8&csf=1&web=1&e=6m2f5h)  [C0](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/credits_iter_0.csv?d=w7c33fa7f879c4168897e76cbccf97195&csf=1&web=1&e=fC1Oj5) |  | **Q17** The number of movies with the following characteristics is \_\_\_\_\_\_\_:  - Spent more than 1,000,000  - The runtime is above 100 minutes or below 60 minutes  - The original language is abbreviated as 'fr' or 'es' |  | 23 | [M1](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/movies_iter_1.csv?d=w2f4610778fbd4a93a3bc2fa1826d7ab0&csf=1&web=1&e=l6IkWp)  [C1](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/credits_iter_1.csv?d=w3a3aea5d9b9343ba846a3df577db3852&csf=1&web=1&e=zj2B74) | **Q17** The number of movies with the following characteristics is \_\_\_\_\_\_\_:  - Spent more than 500,000  - The main genre is Animation  - The runtime is above 120 minutes or below 110 minutes  - The original language is abbreviated as 'fr' or 'es' |  | 4 | [M2](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/movies_iter_2.csv?d=w59cf9b4777884079a73ef5e33b47b6e5&csf=1&web=1&e=pQ9SiD)  [C2](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/credits_iter_2.csv?d=wd4b699f804ad4aa3adf5604517ec904e&csf=1&web=1&e=tMGGNb) | **Q17** The number of movies with the following characteristics is \_\_\_\_\_\_\_:  - Spent more than 5,000,000  -The main genre is Fantasy  - Thevote\_average is above 7 or below 5  - IThe original language is abbreviated as 'en' or 'es' |  | 19 | [M3](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/movies_iter_3.csv?d=wafc95aab49dd477d8f87d52120d14004&csf=1&web=1&e=ehvJeq)  [C3](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/credits_iter_3.csv?d=w5c21e77f92e240dd879cc0066fcdd3c0&csf=1&web=1&e=O6ZKnv) | **Q17** The number of movies with the following characteristics is \_\_\_\_\_\_\_:  - Spent more than 500,000  - The main genre is Crime  - The vote\_average is above 7 or below 5  - The original language is abbreviated as 'ko' or 'es' |  | 2 | [M4](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/movies_iter_4.csv?d=w8eb11c0ab8c7410d927e141ed6fcf940&csf=1&web=1&e=7nR2Tc)  [C4](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/credits_iter_4.csv?d=w137f1d92acf848ef806753ce2c88561e&csf=1&web=1&e=zR40Fl) |
| 10 | DS1-2 | All that apply | **Q18** Which plot would you use to study the distribution of the budget variable? | A)scatter plot  B) boxplot  C) histogram  D) pie chart | B, C | [M0](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/movies_iter_0.csv?d=w3a1f4b08fce7466e851ecde5ee5f0eb8&csf=1&web=1&e=6m2f5h)  [C0](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/credits_iter_0.csv?d=w7c33fa7f879c4168897e76cbccf97195&csf=1&web=1&e=fC1Oj5) |  | **Q18** Which plot would you use to study the distribution of the original\_language variable? | A) scatter plot  B) pie chart  C) heatmap | B | [M1](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/movies_iter_1.csv?d=w2f4610778fbd4a93a3bc2fa1826d7ab0&csf=1&web=1&e=l6IkWp)  [C1](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/credits_iter_1.csv?d=w3a3aea5d9b9343ba846a3df577db3852&csf=1&web=1&e=zj2B74) | **Q18** Which plot would you use to study the relationship between budget and revenue? | A) scatter plot  B) boxplot  C) pie chart  D) histogram | A | [M2](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/movies_iter_2.csv?d=w59cf9b4777884079a73ef5e33b47b6e5&csf=1&web=1&e=pQ9SiD)  [C2](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/credits_iter_2.csv?d=wd4b699f804ad4aa3adf5604517ec904e&csf=1&web=1&e=tMGGNb) | **Q18** Which plot would you use to study the distribution of vote\_average? | A) scatter plot  B) boxplot  C) pie chart  D) histogram | B, D | [M3](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/movies_iter_3.csv?d=wafc95aab49dd477d8f87d52120d14004&csf=1&web=1&e=ehvJeq)  [C3](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/credits_iter_3.csv?d=w5c21e77f92e240dd879cc0066fcdd3c0&csf=1&web=1&e=O6ZKnv) | **Q18** Which plot would you use to study the distribution of vote\_count? | A) scatter plot  B) boxplot  C) pie chart | B | [M4](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/movies_iter_4.csv?d=w8eb11c0ab8c7410d927e141ed6fcf940&csf=1&web=1&e=7nR2Tc)  [C4](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/credits_iter_4.csv?d=w137f1d92acf848ef806753ce2c88561e&csf=1&web=1&e=zR40Fl) |
| 28 | DS1-5 | Fill in the blanks | **Q19** Combine the movie and credits datasets based on a shared attribute. Fill in the missing values  combined\_df = movies.\_\_\_\_\_\_\_(credits, left\_on = \_\_\_\_\_\_\_, right\_on = \_\_\_\_\_\_\_, how = ‘left’)  **Continue your analysis using the combined\_df**  Below you can find a description of each dataset’s columns. **The Credit dataset contains the following features:**  **movie\_id** - A unique identifier for each movie.  **Title -** The movie title  **cast** - The name of lead and supporting actors.  **crew** - The name of Director, Editor, Composer, Writer etc.  **The Movie dataset has the following features:**  **budget** - The budget in which the movie was made.  **genres** - The genres of the movie, Action, Comedy ,Thriller etc.  **homepage** - A link to the homepage of the movie.  **id** - This is in fact the movie\_id as in the first dataset.  **keywords** - The keywords or tags related to the movie.  **original\_language** - The language in which the movie was made.  **original\_title** - The title of the movie before translation or adaptation.  **overview** - A brief description of the movie.  **popularity** - A numeric quantity specifying the movie popularity.  **production\_companies** - The production house of the movie.  **production\_countries** - The country in which it was produced.  **release\_date** - The date on which it was released.  **revenue** - The worldwide revenue generated by the movie.  **runtime** - The running time of the movie in minutes.  **spoken\_languages** - languages in the movie  **status** - "Released" or "Rumored".  **tagline** - Movie's tagline.  **title** - Title of the movie.  **vote\_average** - average ratings the movie received.  **vote\_count** - the count of votes received. |  | merge, id, movie\_id | [M0](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/movies_iter_0.csv?d=w3a1f4b08fce7466e851ecde5ee5f0eb8&csf=1&web=1&e=6m2f5h)  [C0](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/credits_iter_0.csv?d=w7c33fa7f879c4168897e76cbccf97195&csf=1&web=1&e=fC1Oj5) |  | **Q19**Combine the movie and credits datasets based on a shared attribute. Fill in the missing values  combined\_df = movies. \_\_\_\_\_\_\_(credits, left\_on = \_\_\_\_\_\_\_, right\_on = \_\_\_\_\_\_\_, how = ‘left’)  **Continue your analysis using the combined\_df**  Below you can find a description of each dataset’s columns. **The Credit dataset contains the following features:**  **movie\_id** - A unique identifier for each movie.  **Title -** The movie title  **cast** - The name of lead and supporting actors.  **crew** - The name of Director, Editor, Composer, Writer etc.  **The Movie dataset has the following features:**  **budget** - The budget in which the movie was made.  **genres** - The genres of the movie, Action, Comedy ,Thriller etc.  **homepage** - A link to the homepage of the movie.  **id** - This is in fact the movie\_id as in the first dataset.  **keywords** - The keywords or tags related to the movie.  **original\_language** - The language in which the movie was made.  **original\_title** - The title of the movie before translation or adaptation.  **overview** - A brief description of the movie.  **popularity** - A numeric quantity specifying the movie popularity.  **production\_companies** - The production house of the movie.  **production\_countries** - The country in which it was produced.  **release\_date** - The date on which it was released.  **revenue** - The worldwide revenue generated by the movie.  **runtime** - The running time of the movie in minutes.  **spoken\_languages** - languages in the movie  **status** - "Released" or "Rumored".  **tagline** - Movie's tagline.  **title** - Title of the movie.  **vote\_average** - average ratings the movie received.  **vote\_count** - the count of votes received. |  | merge, id, movie\_id | [M1](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/movies_iter_1.csv?d=w2f4610778fbd4a93a3bc2fa1826d7ab0&csf=1&web=1&e=l6IkWp)  [C1](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/credits_iter_1.csv?d=w3a3aea5d9b9343ba846a3df577db3852&csf=1&web=1&e=zj2B74) | **Q19**Combine the movie and credits datasets based on a shared attribute. Fill in the missing values  combined\_df = movies. \_\_\_\_\_\_\_(credits, left\_on = \_\_\_\_\_\_\_, right\_on = \_\_\_\_\_\_\_, how = ‘left’)  **Continue your analysis using the combined\_df**  Below you can find a description of each dataset’s columns. **The Credit dataset contains the following features:**  **movie\_id** - A unique identifier for each movie.  **Title -** The movie title  **cast** - The name of lead and supporting actors.  **crew** - The name of Director, Editor, Composer, Writer etc.  **The Movie dataset has the following features:**  **budget** - The budget in which the movie was made.  **genres** - The genres of the movie, Action, Comedy ,Thriller etc.  **homepage** - A link to the homepage of the movie.  **id** - This is in fact the movie\_id as in the first dataset.  **keywords** - The keywords or tags related to the movie.  **original\_language** - The language in which the movie was made.  **original\_title** - The title of the movie before translation or adaptation.  **overview** - A brief description of the movie.  **popularity** - A numeric quantity specifying the movie popularity.  **production\_companies** - The production house of the movie.  **production\_countries** - The country in which it was produced.  **release\_date** - The date on which it was released.  **revenue** - The worldwide revenue generated by the movie.  **runtime** - The running time of the movie in minutes.  **spoken\_languages** - languages in the movie  **status** - "Released" or "Rumored".  **tagline** - Movie's tagline.  **title** - Title of the movie.  **vote\_average** - average ratings the movie received.  **vote\_count** - the count of votes received. |  | merge, id, movie\_id | [M1](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/movies_iter_1.csv?d=w2f4610778fbd4a93a3bc2fa1826d7ab0&csf=1&web=1&e=l6IkWp)  [C1](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/credits_iter_1.csv?d=w3a3aea5d9b9343ba846a3df577db3852&csf=1&web=1&e=zj2B74) | **Q19** Combine the movie and credits datasets based on a shared attribute. Fill in the missing values  combined\_df = movies. \_\_\_\_\_\_\_(credits, left\_on = \_\_\_\_\_\_\_, right\_on = \_\_\_\_\_\_\_, how = ‘left’)  **Continue your analysis using the combined\_df**  Below you can find a description of each dataset’s columns. **The Credit dataset contains the following features:**  **movie\_id** - A unique identifier for each movie.  **Title -** The movie title  **cast** - The name of lead and supporting actors.  **crew** - The name of Director, Editor, Composer, Writer etc.  **The Movie dataset has the following features:**  **budget** - The budget in which the movie was made.  **genres** - The genres of the movie, Action, Comedy ,Thriller etc.  **homepage** - A link to the homepage of the movie.  **id** - This is in fact the movie\_id as in the first dataset.  **keywords** - The keywords or tags related to the movie.  **original\_language** - The language in which the movie was made.  **original\_title** - The title of the movie before translation or adaptation.  **overview** - A brief description of the movie.  **popularity** - A numeric quantity specifying the movie popularity.  **production\_companies** - The production house of the movie.  **production\_countries** - The country in which it was produced.  **release\_date** - The date on which it was released.  **revenue** - The worldwide revenue generated by the movie.  **runtime** - The running time of the movie in minutes.  **spoken\_languages** - languages in the movie  **status** - "Released" or "Rumored".  **tagline** - Movie's tagline.  **title** - Title of the movie.  **vote\_average** - average ratings the movie received.  **vote\_count** - the count of votes received. |  | merge, id, movie\_id | [M1](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/movies_iter_1.csv?d=w2f4610778fbd4a93a3bc2fa1826d7ab0&csf=1&web=1&e=l6IkWp)  [C1](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/credits_iter_1.csv?d=w3a3aea5d9b9343ba846a3df577db3852&csf=1&web=1&e=zj2B74) | **Q19** Combine the movie and credits datasets based on a shared attribute. Fill in the missing values  combined\_df = movies. \_\_\_\_\_\_\_(credits, left\_on = \_\_\_\_\_\_\_, right\_on = \_\_\_\_\_\_\_, how = ‘left’)  **Continue your analysis using the combined\_df**  Below you can find a description of each dataset’s columns. **The Credit dataset contains the following features:**  **movie\_id** - A unique identifier for each movie.  **Title -** The movie title  **cast** - The name of lead and supporting actors.  **crew** - The name of Director, Editor, Composer, Writer etc.  **The Movie dataset has the following features:**  **budget** - The budget in which the movie was made.  **genres** - The genres of the movie, Action, Comedy ,Thriller etc.  **homepage** - A link to the homepage of the movie.  **id** - This is in fact the movie\_id as in the first dataset.  **keywords** - The keywords or tags related to the movie.  **original\_language** - The language in which the movie was made.  **original\_title** - The title of the movie before translation or adaptation.  **overview** - A brief description of the movie.  **popularity** - A numeric quantity specifying the movie popularity.  **production\_companies** - The production house of the movie.  **production\_countries** - The country in which it was produced.  **release\_date** - The date on which it was released.  **revenue** - The worldwide revenue generated by the movie.  **runtime** - The running time of the movie in minutes.  **spoken\_languages** - languages in the movie  **status** - "Released" or "Rumored".  **tagline** - Movie's tagline.  **title** - Title of the movie.  **vote\_average** - average ratings the movie received.  **vote\_count** - the count of votes received. |  | merge, id, movie\_id | [M1](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/movies_iter_1.csv?d=w2f4610778fbd4a93a3bc2fa1826d7ab0&csf=1&web=1&e=l6IkWp)  [C1](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/credits_iter_1.csv?d=w3a3aea5d9b9343ba846a3df577db3852&csf=1&web=1&e=zj2B74) |
| 37 | DS1-7 | Multiple choice | **Q20** Find the total number of columns after combining the dataframes and by removing redundant columns. | A)22  B) 23  C) 21  D) 24 | A | [M0](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/movies_iter_0.csv?d=w3a1f4b08fce7466e851ecde5ee5f0eb8&csf=1&web=1&e=6m2f5h)  [C0](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/credits_iter_0.csv?d=w7c33fa7f879c4168897e76cbccf97195&csf=1&web=1&e=fC1Oj5) |  | **Q20** Find the total number of columns after combining the dataframes and by removing redundant columns. | A)22  B) 23  C) 21  D) 24 | B | [M1](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/movies_iter_1.csv?d=w2f4610778fbd4a93a3bc2fa1826d7ab0&csf=1&web=1&e=l6IkWp)  [C1](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/credits_iter_1.csv?d=w3a3aea5d9b9343ba846a3df577db3852&csf=1&web=1&e=zj2B74) | **Q20** Find the total number of columns after combining the dataframes and by removing redundant columns. | A)22  B) 23  C) 21  D) 24 | D | [M2](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/movies_iter_2.csv?d=w59cf9b4777884079a73ef5e33b47b6e5&csf=1&web=1&e=pQ9SiD)  [C2](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/credits_iter_2.csv?d=wd4b699f804ad4aa3adf5604517ec904e&csf=1&web=1&e=tMGGNb) | **Q20** Find the total number of rows after combining the dataframes and removing the rows that did not get any information about credits.  Remove these records from the analysis | A)4489  B)4480  C)4450  D)4459 | B | [M3](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/movies_iter_3.csv?d=wafc95aab49dd477d8f87d52120d14004&csf=1&web=1&e=ehvJeq)  [C3](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/credits_iter_3.csv?d=w5c21e77f92e240dd879cc0066fcdd3c0&csf=1&web=1&e=O6ZKnv) | **Q20** Find the total number of rows after combining the dataframes and removing the rows that did not get any information about credits.  Remove these records from the analysis | A)4489  B)4480  C)4357  D)4365 | C | [M4](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/movies_iter_4.csv?d=w8eb11c0ab8c7410d927e141ed6fcf940&csf=1&web=1&e=7nR2Tc)  [C4](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/credits_iter_4.csv?d=w137f1d92acf848ef806753ce2c88561e&csf=1&web=1&e=zR40Fl) |
| 22 | DS1-4 | Multiple choice | **Q21** Calculate the weighted rating (wr) for the movie  WeightedRating(WR)=[(v.R)/(v+m)] + [(m.C)/(v+m)]  here,  **v** is the number of votes for the movie  **m** is the minimum votes required to be listed in the chart  **R** is the vote\_average of the movie  **C** is the mean vote\_average across the whole report  For calculation of m, we will use 90th percentile as our cutoff. In other words, for a movie to feature in the charts, it must have more votes than at least 90% of the movies in the list.    List the five movies with the highest WR. | A) Inception, Forrest Gump, The Lord of the Rings: The Fellowship of the Ring, The Empire Strikes Back, Star Wars  B) The Shawshank Redemption, The Godfather, Interstellar, The Lord of the Rings: The Return of the King, The Empire Strikes Back  C) The Godfather, Inception, Forrest Gump, Interstellar, The Lord of the Rings: The Return of the King  D) Fight Club, Pulp Fiction, The Godfather, Inception, The Empire Strikes Back | B | [M0](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/movies_iter_0.csv?d=w3a1f4b08fce7466e851ecde5ee5f0eb8&csf=1&web=1&e=6m2f5h)  [C0](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/credits_iter_0.csv?d=w7c33fa7f879c4168897e76cbccf97195&csf=1&web=1&e=fC1Oj5) |  | **Q21** Calculate the weighted rating (wr) for the movie  WeightedRating(WR)=[(v.R)/(v+m)] + [(m.C)/(v+m)]  where,  **v** is the number of votes for the movie  **m** is the minimum votes required to be listed in the chart  **R** is the vote\_average of the movie  **C** is the mean vote\_average across the whole report  For calculation of m, we will use 95th percentile as our cutoff. In other words, for a movie to feature in the charts, it must have more votes than at least 95% of the movies in the list.    List the five movies with the highest WR. | A)Inception, Forrest Gump, The Lord of the Rings: The Fellowship of the Ring, The Empire Strikes Back, Star Wars    B) Fight Club, Pulp Fiction, The Godfather, Inception, The Empire Strikes Back    C) Inception, Forrest Gump, The Empire Strikes Back, Schindler's List, Whiplash    D) The Godfather, Inception, Forrest Gump, Interstellar, The Lord of the Rings: The Return of the King | A | [M1](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/movies_iter_1.csv?d=w2f4610778fbd4a93a3bc2fa1826d7ab0&csf=1&web=1&e=l6IkWp)  [C1](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/credits_iter_1.csv?d=w3a3aea5d9b9343ba846a3df577db3852&csf=1&web=1&e=zj2B74) | **Q21** Calculate the weighted rating (wr) for the movie  WeightedRating(WR)=[(v.R)/(v+m)] + [(m.C)/(v+m)]  where,  **v** is the number of votes for the movie  **m** is the minimum votes required to be listed in the chart  **R** is the vote\_average of the movie  **C** is the mean vote\_average across the whole report  For calculation of m, we will use 85th percentile as our cutoff. In other words, for a movie to feature in the charts, it must have more votes than at least 85% of the movies in the list.    List the five movies with the highest WR. | A) Inception, Forrest Gump, The Lord of the Rings: The Fellowship of the Ring, The Empire Strikes Back, Star Wars  B) The Shawshank Redemption, The Godfather, Interstellar, The Lord of the Rings: The Return of the King, The Empire Strikes Back  C) Fight Club, Pulp Fiction, The Godfather, Inception, The Empire Strikes Back  D) The Godfather, Inception, Forrest Gump, Interstellar, The Lord of the Rings: The Return of the King | C | [M2](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/movies_iter_2.csv?d=w59cf9b4777884079a73ef5e33b47b6e5&csf=1&web=1&e=pQ9SiD)  [C2](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/credits_iter_2.csv?d=wd4b699f804ad4aa3adf5604517ec904e&csf=1&web=1&e=tMGGNb) | **Q21** Calculate the weighted rating (wr) for the movie  WeightedRating(WR)=[(v.R)/(v+m)] + [(m.C)/(v+m)]  where,  **v** is the number of votes for the movie  **m** is the minimum votes required to be listed in the chart  **R** is the vote\_average of the movie  **C** is the mean vote\_average across the whole report  For calculation of m, we will use 85th percentile as our cutoff. In other words, for a movie to feature in the charts, it must have more votes than at least 85% of the movies in the list.    List the five movies with the highest WR. | A) Inception, Forrest Gump, The Empire Strikes Back, Schindler's List, Whiplash  B)Inception, Forrest Gump, The Lord of the Rings: The Fellowship of the Ring, The Empire Strikes Back, Star Wars  C) The Shawshank Redemption, The Godfather, Interstellar, The Lord of the Rings: The Return of the King, The Empire Strikes Back  D) The Godfather, Inception, Forrest Gump, Interstellar, The Lord of the Rings: The Return of the King | A | [M3](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/movies_iter_3.csv?d=wafc95aab49dd477d8f87d52120d14004&csf=1&web=1&e=ehvJeq)  [C3](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/credits_iter_3.csv?d=w5c21e77f92e240dd879cc0066fcdd3c0&csf=1&web=1&e=O6ZKnv) | **Q21** Calculate the weighted rating (wr) for the movie  WeightedRating(WR)=[(v.R)/(v+m)] + [(m.C)/(v+m)]  where,  **v** is the number of votes for the movie  **m** is the minimum votes required to be listed in the chart  **R** is the vote\_average of the movie  **C** is the mean vote\_average across the whole report  For calculation of m, we will use 85th percentile as our cutoff. In other words, for a movie to feature in the charts, it must have more votes than at least 85% of the movies in the list.    List the five movies with the highest WR. | A) Inception, Forrest Gump, The Empire Strikes Back, Schindler's List, Whiplash  B) The Godfather, Inception, Forrest Gump, Interstellar, The Lord of the Rings: The Return of the King  C) Fight Club, Pulp Fiction, The Godfather, Inception, The Empire Strikes Back  D) Inception, Forrest Gump, The Lord of the Rings: The Fellowship of the Ring, The Empire Strikes Back, Star Wars | B) | [M4](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/movies_iter_4.csv?d=w8eb11c0ab8c7410d927e141ed6fcf940&csf=1&web=1&e=7nR2Tc)  [C4](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/credits_iter_4.csv?d=w137f1d92acf848ef806753ce2c88561e&csf=1&web=1&e=zR40Fl) |
| 12 | DS1-3 | All that apply | **Q22** Create the correlation matrix of the dataset and select the appropriate option: | A)Vote count has the maximum correlation coefficient with popularity and revenue  B)Strength of relationship for vote average and id is similar to vote average and popularity  C)A correlation coefficient of 0.6 indicates a stronger relationship than a correlation coefficient of –0.7  D)If two variables have a low pearson correlation coefficient they can’t have any relationship with each other | A | [M0](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/movies_iter_0.csv?d=w3a1f4b08fce7466e851ecde5ee5f0eb8&csf=1&web=1&e=6m2f5h)  [C0](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/credits_iter_0.csv?d=w7c33fa7f879c4168897e76cbccf97195&csf=1&web=1&e=fC1Oj5) |  | **Q22** Create the correlation matrix of the dataset and select the appropriate option: | A) Vote count has the maximum correlation coefficient with popularity and budget    B) A correlation coefficient of 0.6 indicates a stronger relationship than a correlation coefficient of –0.7    C) According to the correlation matrix, it is worth exploring the increase of budget to increase revenue | C | [M1](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/movies_iter_1.csv?d=w2f4610778fbd4a93a3bc2fa1826d7ab0&csf=1&web=1&e=l6IkWp)  [C1](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/credits_iter_1.csv?d=w3a3aea5d9b9343ba846a3df577db3852&csf=1&web=1&e=zj2B74) | **Q22** Create the correlation matrix of the dataset and select the appropriate option: | A) Weighted rating (wr) is positive and highly correlated with budget  B) We should not be studying the correlation between id and the rest of the movies’ attributes  C) It seems promising to increase the budget in order to increase the score | B | [M2](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/movies_iter_2.csv?d=w59cf9b4777884079a73ef5e33b47b6e5&csf=1&web=1&e=pQ9SiD)  [C2](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/credits_iter_2.csv?d=wd4b699f804ad4aa3adf5604517ec904e&csf=1&web=1&e=tMGGNb) | **Q22** Create the correlation matrix of the dataset and select the appropriate option: | A) Score is positive and highly correlated with budget  B) There is a linear relationship between vote\_count and revenue  C) A correlation of -0.8 does not imply a relationship between two variables | B | [M3](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/movies_iter_3.csv?d=wafc95aab49dd477d8f87d52120d14004&csf=1&web=1&e=ehvJeq)  [C3](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/credits_iter_3.csv?d=w5c21e77f92e240dd879cc0066fcdd3c0&csf=1&web=1&e=O6ZKnv) | **Q22** Create the correlation matrix of the dataset and select the appropriate option: | A) There is a linear relashionship between popularity and runtime  B) Score is positive and highly correlated with budget  C) There is a linear relationship between vote\_count and popularity | A | [M4](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/movies_iter_4.csv?d=w8eb11c0ab8c7410d927e141ed6fcf940&csf=1&web=1&e=7nR2Tc)  [C4](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/credits_iter_4.csv?d=w137f1d92acf848ef806753ce2c88561e&csf=1&web=1&e=zR40Fl) |
| 29 | DS1-5 | Fill in the blank | **Q23** What was the budget (integer) of the film with a crew of 150 |  | 78000000 | [M0](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/movies_iter_0.csv?d=w3a1f4b08fce7466e851ecde5ee5f0eb8&csf=1&web=1&e=6m2f5h)  [C0](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/credits_iter_0.csv?d=w7c33fa7f879c4168897e76cbccf97195&csf=1&web=1&e=fC1Oj5) |  | **Q23** What was the vote average of the film with a crew of 168? |  | 6.9 | [M1](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/movies_iter_1.csv?d=w2f4610778fbd4a93a3bc2fa1826d7ab0&csf=1&web=1&e=l6IkWp)  [C1](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/credits_iter_1.csv?d=w3a3aea5d9b9343ba846a3df577db3852&csf=1&web=1&e=zj2B74) | **Q23** What was the budget (integer) of the film with a crew of 151 |  | 85000000 | [M2](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/movies_iter_2.csv?d=w59cf9b4777884079a73ef5e33b47b6e5&csf=1&web=1&e=pQ9SiD)  [C2](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/credits_iter_2.csv?d=wd4b699f804ad4aa3adf5604517ec904e&csf=1&web=1&e=tMGGNb) | **Q23** What was the budget (integer) of the film with a crew of 130 |  | 30000000 | [M3](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/movies_iter_3.csv?d=wafc95aab49dd477d8f87d52120d14004&csf=1&web=1&e=ehvJeq)  [C3](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/credits_iter_3.csv?d=w5c21e77f92e240dd879cc0066fcdd3c0&csf=1&web=1&e=O6ZKnv) | **Q23** What was the vote\_average of the film with a crew of 160 |  | 7.2 | [M4](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/movies_iter_4.csv?d=w8eb11c0ab8c7410d927e141ed6fcf940&csf=1&web=1&e=7nR2Tc)  [C4](https://worldbankgroup.sharepoint.com/:x:/r/teams/DevelopmentDataPartnershipCommunity-WBGroup/Shared%20Documents/Projects/Data%20Lab/Certifications/data_science_skills/datasets/credits_iter_4.csv?d=w137f1d92acf848ef806753ce2c88561e&csf=1&web=1&e=zR40Fl) |
| 41  same question across all iterations | DS1-8 | All that apply | **Q24** Which of the following file formats can be used to save the merged dataframe | A) png  B) tiff  C) csv  D) xlsx  E) doc | C, D |  |  | **Q24** Which of the following file formats can be used to save the merged dataframe | A) png  B) tiff  C) csv  D) xlsx  E) doc | C, D |  | **Q24**Which of the following file formats can be used to save the merged dataframe | A) png  B) tiff  C) csv  D) xlsx  E) doc | C, D |  | **Q24** Which of the following file formats can be used to save the merged dataframe | A) png  B) tiff  C) csv  D) xlsx  E) doc | C, D |  | **Q24** Which of the following file formats can be used to save the merged dataframe | A) png  B) tiff  C) csv  D) xlsx  E) doc | C, D |  |
| 42 |  |  | Please upload your notebook.script with the code you had written to obtain the solutions. The code will be reviewed by the evaluator who will determine whether you pass the exam or not. |  |  |  |  | Please upload your notebook.script with the code you had written to obtain the solutions. The code will be reviewed by the evaluator who will determine whether you pass the exam or not. |  |  |  | Please upload your notebook.script with the code you had written to obtain the solutions. The code will be reviewed by the evaluator who will determine whether you pass the exam or not. |  |  |  | Please upload your notebook.script with the code you had written to obtain the solutions. The code will be reviewed by the evaluator who will determine whether you pass the exam or not. |  |  |  | Please upload your notebook.script with the code you had written to obtain the solutions. The code will be reviewed by the evaluator who will determine whether you pass the exam or not. |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |