**Раздел 01.Математика и логика**

|  |  |
| --- | --- |
| Вопрос | Ответы |
| Известно, что один секретарь набирает 100-страничный документ за 2 часа, а другой - 3 часа. Сколько часов потребуется им, чтобы набрать такого же объема документ работая над ним совместно? | ***1,2***  2,5  1,5  1 |
| Предположим, что у Вас есть данные по 10 клиентам. Вам нужно попарно сравнить показатели по этим клиентам. Сколько получится таких уникальных попарных сравнений? | 10  ***45***  90  25 |
| Предположим, что выражение «Только Юридические лица совершают переводы» - истинно. Какой из приведенных выводов можно сделать из этого утверждения? | ***a. Все физические лица не совершают переводы***  b. Любой, кто не пользуется переводами, является физическим лицом  c. Некоторые из физических лиц пользуются переводами  d. Только физические лица не пользуются переводами  е. Юридические лица не пользуются ничем кроме переводов |
| В компании есть три отдела A,B,C которые договорились о совместном участии в проектах , а именно а) если отдел В не участвует в проекте, то в этом проекте не участвует и отдел А; в) если отдел В принимает участие в проекте, то в нем принимают участие отделы А и С. Какое из перечисленных утверждений является верным? | А. Отделы А и С никогда не участвуют в проектах вместе  В. Отдел А может работать в проектах в одиночку  ***С. Отдел С должен участвовать в проекте, если в нем участвует отдел А***  D. Отдел С не может работать над проектом в одиночку |
| Два аналитика с разной скоростью работают над отчетами. Первый аналитик делает 15 отчетов за рабочий день, а второй – 10. Сколько рабочих дней им потребуется на то, чтобы выполнить план отдела и сделать 100 отчетов? | А. 5  ***В. 4***  С. 10  D.12 |
| В компании есть дата инженеры, дата аналитики и дата сайентисты. Известно следующее: а) не все дата инженеры являются дата аналитиками; б) если дата сайентист не является дата аналитиком, то он не дата инженер; Какой из приведенных выводов можно сделать из этих утверждений? | А. Все дата инженеры - дата сайентисты  В. Все дата аналитики – дата инжененры  ***С. Не все дата инженеры - дата сайентисты***  D.Все дата сайентисты – дата инженеры |
| Предположим, что ваш отдел состоит из 5 человек. Вы переехали в новую комнату. В ней есть 6 рабочих мест. Сколькими способами можно расположить сотрудников отдела в новой комнате? | ***А. 120***  В. 720  С. 6  D. 30 |

**Раздел 02 Основы эконометрики**

|  |  |
| --- | --- |
| Даны значения ряда 12,13,22,12,49,10,13,25,12. Посчитайте моду ряда | ***А. 12***  В. 39  С. 19  D.13 |
| Коэффициент корреляции, равный 1, говорит о том, что | А. Между переменными нет никакой зависимости  В. Между переменными нет линейной зависимости  ***С. Между переменными существует функциональная зависимость***  D.Между переменными есть статистическая зависимость |
| Что может означать, если парный коэффициент корреляции между признаками больше частного коэффициента корреляции? | А. другие признаки уменьшают взаимосвязь между изучаемыми признаками  ***В. Между изучаемыми признаками существует функциональная зависимость***  С. Другие признаки усиливают взаимосвязь между изучаемыми признаками  D. это невозможно |
| Даны значения ряда 12,15,7,8,22. Посчитайте размах вариации | А. 12  В. 12,8  ***С. 15***  D. 6 |
| Даны значения ряда 12,13,22,12,49,10,13,25,12. Посчитайте медиану | А. 14  В. 12  С. 19  ***D.13*** |
| Перед вами уравнение линейной регрессии: y=2x\_1-1,5x\_2+10x\_3. Выберите верное утверждение о нем. | ***А. при увеличении х1 на единицу, у увеличивается на 2 единицы***  В. При х2 и х3 равным нулю, у тоже равен нулю  С. При уменьшении х3 на единицу, у увеличивается на 10 единиц  D. при увеличении х1 на единицу, у увеличивается в 2 раза |
| Вы построили уравнение линейной регрессии и получили R2 равным 0,64 . как это интерпретировать? | А. 64% наблюдений получили верный прогноз по модели  В. При увеличении значений всех признаков на одну единицу, у увеличивается на 0,64  ***С. Модель объясняет 64% дисперсии целевой переменной***  D. коэффициент корреляции между признаками и целевой переменной равен 0,64 |

**Раздел 03. Финансовый и инвестиционный анализ**

|  |  |
| --- | --- |
| Каков срок окупаемости проекта, если изначальные инвестиции составили 1 000 000 руб., а дисконтированный денежный поток в течении 6 лет выглядит так: -200 000, -50 000, +100 000, +300 000, +300 000, +250 000? | А. 5 лет  В. 6 лет  ***С. Проект не успевает окупиться за эти 6 лет***  D. 4 года |
| Компания положила 720 000 руб. на депозит сроком на 3 года с условием, что общая сумма возврата составит 890 000 руб. Какова процентная ставка в год для случая сложного процента? | А. 19,1%  В. 41,2%  С. 23,6%  ***D. 7,3%*** |
| Согласно оценке, проект будет генерировать чистую прибыль в размере 200 000 руб. в течение 5 лет. Ставка дисконтирования для проекта – 20%. Чему равен NPV? | А. 1 000 000 руб.  В. 1 641 41 руб.  ***С. 598 120 руб.***  D. 717 750 руб. |
| Что из перечисленного обладает наименьшей ликвидностью? | А. Деньги на депозитах  ***В. Деньги, вложенные в недвижимость***  С. Деньги на текущих счетах  D. наличные деньги |
| Человек взял в кредит 500 000 руб. по ставке 12% годовых (простой процент) Через какой период времени согласно этим условиям, он вынужден будет вернуть в банк 1 000 000 руб.? | А. 6 лет и 1 мес.  ***В. 8 лет и 4 мес.***  С. 7 лет и 4 мес.  D. 10 лет |
| Оценивая инвестиционную привлекательность проекта, вы рассчитали внутреннюю норму доходности (IRR). Она оказалась меньше ставки дисконтирования, которую вы закладывали в финансовую модель. Какой вывод можно сделать? | А. инвестиции в этот проект целесообразны, только проект будет долго окупаться  ***В. Инвестиции в этот проект целесообразны, проект быстро окупится***  С. Связи между внутренней нормой доходности и ставкой дисконтирования не существует  D.инвестиции в этот проект нецелесообразны, проект никогда не окупится |
| Для чего при анализе проекта денежный поток дисконтируется? | ***А. деньги, полученные сейчас, ценнее, чем деньги, полученные потом***  В. Деньги с течением времени обесцениваются за счет инфляции  С. Деньги нужно привлекать, и у них есть своя стоимость - дисконт  D.деньги, полученные потом, ценнее , чем деньги , полученные сейчас |

**Раздел 04. Анализ данных и машинное обучение**

|  |  |
| --- | --- |
| К какому типу задач машинного обучения относится задача определения вероятности того, что человек вернет кредит? | ***А. классификация***  В. кластеризация  С. регрессия  D. снижения размерности |
| Проанализировав график ниже, подсчитайте чему приблизительно равна медиана? | А. 17  В. 44  С. 29  ***D.24*** |
| Проанализировав график ниже, подсчитайте чему приблизительно равен интерквартильный размах | А. 7  В. 35  ***С. 12***  D.5 |
| Можно ли считать модель на графике ниже (синяя линяя) качественной | А. нет модель плохо предсказывает наблюдения обучающей выборки  ***В. Да , модель идеально предсказывает наблюдения обучающей выборки, проходя через все точки***  С. Нет, поскольку она не является линейкой  D. нет, модель переобучилась |
| С какой целью производится нормирование данных? | А. для улучшения качества линейных моделей  В. Для борьбы с мультиколлинеарностью  С. Для улучшения интерпретируемости моделей  ***D.для борьбы с аномальными наблюдениями*** |
| Если поделить количество правильно угаданных классов на общее число наблюдений, то мы получим метрику качества: | А. RMSE  В. ROC-AUC  С. precision  ***D.accuracy*** |
| Если у признака 99% значений пропущены, то какой из перечисленных способов работы с пропусками лучше всего проходит | ***А. удалить признак целиком***  В. Заполнить пропущенные значения средним  С. Заполнить пропущенные значения медианой  D. Заполнить пропущенные значения, используя уравнения регрессии |

**Раздел 05. Интерпретация данный и оценка значимости результатов**

|  |  |
| --- | --- |
| В группе 1 после коммуникации продукт оформили 1,5% клиентов, в группе 2 -1% клиентов. Можно ли сказать, что коммуникация в группе 1 сработала лучше? | А. Да, т.к разница в откликах 0,5%  В. Нет т.к. 0,5% - слишком маленькая разница, чтобы сделать вывод о значимых результатах  ***С. Не хватает данных для точного ответа***  D. Да, если продукт в каждой группе оформили более 100 клиентов  Е. Да, если в исходных группах больше 100 клиентов |
| Для проведения А/В-теста выделяется 2 группы, каждой из которых будут направлены разные SMS-сообщения с предложением оформить продукт. Группе А – обычное старое сообщение, отправлявшееся по таким кампаниям раньше, группе В – новое. Чтобы определить естественный спрос, необходимо: | ***А. заранее выделить группу клиентов, подобную одной из групп выше, которой коммуникации отправляться не будут, и рассчитать по ней % оформлений продукта***  В. Рассчитать % оформлений продукта по всем остальным клиентам, не попавшим в группы А и В и, соответственно, не получавшим коммуникацию  С. Рассчитать отклик по группе А с обычным сообщением - это и есть естественный спрос  D.рассчитать разницу между откликом в группе А и откликом в группе В |
| В течение января планируется обзвонить группу клиентов, отобранных по ряду критериев для прохождения оформить кредитную карту. (здесь и далее Отклик = % оформивших предлагаемый продукт клиентов в рассматриваемой группе клиентов).  Чтобы рассчитать эффект от коммуникации, необходимо: | А. сравнить отклик по клиентам, до которых дозвонились, с откликом по клиентам, с которыми не смогли связаться  В. Сравнить отклик по этой группе клиентов с откликом по остальным клиентам, не попавшим в выборку для проводимой кампании  С. Все ответы не верны  D.сравнить отклик по этой группе клиентов с откликом по ним же месяцем ранее  ***Е. все ответы верны*** |
| Была проведена кампания на 2 сегментах клиентов, для каждого из которых была выделена контрольная группа – результаты по этим сегментам представлены в таблице ниже. Чему будут равны доп.отклики и доп.сделки по всем клиентам, отобранным в оба сегмента? | А. 1,5%, 9 000  ***В. 0,75%, 5 250***  С. 0,75%, 4 500  D.0,875%, 5 250  Е. 1,5%, 5 250 |

**Раздел 06. SQL**

|  |  |
| --- | --- |
| Какой из запросов вернет корректный результат, выдав всех клиентов, открывших счет, начиная с 2010 года? | ***А. SELECT Id, YEAR(AccountDate) AS AccountYear***  ***FROM Clients***  ***WHERE YEAR(AccountDate)>=2010***  В. SELECT Id, YEAR(AccountDate) AS AccountYear  FROM Clients  HAVING YEAR(AccountDate)>=2010  С. SELECT Id, YEAR(AccountDate) AS AccountYear  FROM Clients  HAVING AccountYear>=2010  D. SELECT Id, YEAR(AccountDate) AS AccountYear  FROM Clients  WHERE AccountYear>=2010 |
| Какая из перечисленных конструкций использования оператора having верна? | А. select\*from [таблица] group by [поля] where [условия] having [условия]  В. select\*from [таблица] having [условия] group by [поля] where [условия]  С. select\*from [таблица] group by [поля] having [условия] where [условия]  ***D. select\*from [таблица] where [условия] group by [поля] having [условия]***  Е. select\*from [таблица] where [условия] having [условия] group by [поля] |
| Какое из перечисленных ограничений в команде create table позволяет определить область значений для поля таблицы | ***А. check***  В. default  С. references  D. unique  E. key |
| Какой из запросов позволит найти дубликаты записей с несколькими полями | А. SELECT name, email, COUNT(\*)  FROM users  GROUP BY name, email  WHERE COUNT(\*)>1  ***В. SELECT name, email, COUNT(\*)***  ***FROM users***  ***GROUP BY name, email***  ***HAVING COUNT(\*)>1***  С. SELECT name, COUNT (email)  FROM users  GROUP BY email  HAVING COUNT(email)>1  D. SELECT name, COUNT (email)  FROM users  GROUP BY email  HAVING COUNT(email)>2 |
| Предположим, есть две таблицы First (ID, e-mail,occupation), Second(ID,e-mail). В первой таблице содержится 12 записей, а во второй – 5. Сколько строк будет в результате выполнения следующего запроса: Select\*From First,Second; | ***А. 60***  В. неизвестно  С. 17  D.12 |
| Как можно в таблице My Table заменить значения поля Status из unemployed fired? | А. MODIFY MyTable SET Status = ‘unemployed’ WHERE Status = ‘fired’  В. CHANGE MyTable.Status FROM ‘unemployed’ TO ‘fired’  С. UPDATE MyTable.Status SET ‘fired’ WHERE Status =‘unemployed’  ***D. UPDATE MyTable SET Status = ‘fired’ WHERE Status=‘unemployed’***  E. MODIFY MyTable SET Status = ‘fired’ WHERE Status = ‘unemployed’ |
| Нужно вывести список товаров из таблицы (product), проданных на сумму более 10000 без указания суммы. Сделки хранятся в таблице (order). Какой из перечисленных фрагментов кода позволит решить поставленную задачу? | А. select p.name from product p inner join order o on p. product\_id = o. product\_id group by p.name where sum  В. select p.name from product p inner join order o on p. product\_id = o. product\_id group by p.name having (p.price\*o.quantity)>10000  С. select p.name from product p, order o where p. product\_id = p. product\_id and sum (p.price\*o.quantity)>10000 group by p.name  D. select p.name from product p inner join order o on p. product\_id = o. product\_id having sum(p.price\*o.quantity)>10000 group by p.name  ***E. select p.name from product p inner join order o on p. product\_id = o. product\_id group by p.name having sum(p.price\*o.quantity)>10000*** |
| Предположим, у вас есть поле со значениями 0 и 1. Вам необходимо написать запрос, который заменит 0 на 999, а 1 – на 100. Какой из перечисленных запросов выдаст корректный результат для сформулированной задачи? | А. SELECT\*FROM Features  WHERE Gender = 0  REPLACE 999  ELSE  REPLACE 100  ***В. UPDATE Features SET Gender = CASE WHEN Gender =0 THEN 999 ELSE 100 END***  С. UPDATE Features SET Gender = CASE WHERE IF Gender = 0 THEN 999 ELSE 100 END  D. SELECT\*FROM Features  UPDATE GENDER = CASE WHEN Gender = 0 THEN 999  ELSE 100 |
| Необходимо вычислить средний объем продаж (advance) книг годовой доход (sales) от продаж книг по всем категориям (type) из таблицы titles. Какой из перечисленных фрагментов кода позволит решить поставленную задачу? | ***А. Select avg (advance), sum (sales) from titles group by type***  В. Select avg (advance), sum (sales) from titles  С. Select avg (advance), sum (sales) from titles where type between ‘business’ and ‘internet’  D. Select avg (advance), sum (sales) from titles where type = ‘business’ and type ‘internet’  E. Select avg (advance), sum (sales) from titles where type between ‘business’ or ‘internet’ |
| Укажите верный синтаксис для копирования данных из одной таблицы в другую? | А. INSERT INTO table2(coloumn1, coloumn2, coloumn3,…)  SELECT coloumn1, coloumn2, coloumn3,…  FROM table1  HAVING condition  В. INSERT FROM table1 INTO table2 (coloumn1, coloumn2, coloumn3,…)  SELECT coloumn1, coloumn2, coloumn3,…  FROM table1  WHERE condition  ***С. INSERT INTO table2 (coloumn1, coloumn2, coloumn3,…)***  ***SELECT coloumn1, coloumn2, coloumn3,…***  ***FROM table1***  ***WHERE condition***  D. INSERT INTO table1 INTO table2 (coloumn1, coloumn2, coloumn3,…) FROM table1  SELECT coloumn1, coloumn2, coloumn3,…  WHERE condition |

**Раздел 07. Python**

|  |  |
| --- | --- |
| Что делает оператор break? | A. переходит на «вершину» цикла и запускает следующую итерацию  ***B. выход из текущего выполняющего цикла***  C. выход из программы  D.сбрасывает переменную итерации к ее начальному значению |
| Какие из следующих операторов Python выведут длину списка, хранящегося в переменных данных? | A. print(length(data))  B. print(data.length)  C. print(strlen(data))  D. print(data.length())  ***E. print(len(data))***  F. print(data.Len) |
| Что распечатает следующий код? Подсказка: это вопрос с подвохом, и большинство скажет, что в этом коде есть ошибка – так что читайте внимательней | A. 74  ***B.-1***  C.42  D.3 |
| Чему соответствует «[0-9]+» в регуляторном выражении? | A. ноль или более цифр  B.любое математическое выражение  C.любое количество цифр в начале строки  ***D.одна или несколько цифр***  E. несколько цифр и знак плюс |
| Какой результат будет получен в результате выполнения этого кода: | ***A. Hellothere***  B. Hello there  C. Hello  D. 0 |

**Раздел 08. TERADATA**

|  |  |
| --- | --- |
| Ваша таблица оказалась заблокирована во время вызова MLOAD из-за операции, завершившейся ошибкой. Как ее разблокировать, чтобы снова иметь возможность выполнять операции над ней? | A.UNLOCK<table name>  ***B.RELEASE MLOAD<table name >***  C. RELEASE <table name >  D. UNLOCK MLOAD <table name> |
| В чем отличие TRUMP от MULTILOAD? | A. TRUMP выполняет обновление данных в режиме real-time  B. TRUMP подходит для обработки батчей большого объема  C***. MULTILOAD выполняет обновление данных в режиме real-time***  D. во время выполнения MULTILOAD можно выполнять запросы к таблице |
| Что эффективнее для поиска дубликатов: GROUP BY или DISTINCT? | A. GROUP BY неприменим для поиска дубликатов  B. если дубликатов много, то GROUP BY  C.одинаково  ***D. если дубликатов много, что DISTINCT*** |
| Что из перечисленного является корректным способом множественного добавления данных в таблицу? | A. Insert into Sales “select\*from customer”,Loan”select\*from customer”  B. Insert into Sales “select\*from customer”, Insert into Loan”select\*from customer”  C. Insert into Sales “select\*from customer”, into Loan”select\*from customer”  ***D. Insert into Sales “select\*from customer”,***  ***Insert into Loan”select\*from customer”*** |
| В чем разница между TIMESTAMP(6) и TIMESTAMP(0) | A. TIMESTAMP(0) содержит в себе помимо даты еще и время  ***B. TIMESTAMP(6) содержит в себе миллисекунды***  C. TIMESTAMP(0) содержит в себе миллисекунды  D TIMESTAMP(6) содержит в себе помимо времени и дату |
| Что такое Sparse Index в Teradata? | A. это индекс , в котором пустые значения не хранятся  B.такого индекса не существует  C.это индекс, в котором удалены дубликаты  ***D. это join index, содержащий where*** |
| Предположим, что у вас есть колонка с датами (тип дата). Каким образом из нее можно извлечь месяц? | A. select extract (month from <column name>)  ***B. select (month from <column name>)***  C. использовать substring  D. ничего из перечисленного |
| Какого типа индекса не существует в Teradata? | ***A. foreign index***  B.patitioned primary index  C.secondary index  D.join index  E. hash index  F. primary index |