

SQL Interview Questions

Sample data load:

```
CREATE TABLE `customers` ( `customerNumber` int , `customerName` varchar(50) ,  
`contactLastName` varchar(50) , `contactFirstName` varchar(50) , `phone` varchar(50) ,  
`addressLine1` varchar(50) , `addressLine2` varchar(50) , `city` varchar(50) , `state`  
varchar(50) , `postalCode` varchar(15) , `country` varchar(50) ,  
`salesRepEmployeeNumber` int , `creditLimit` decimal(10,2) ) row format delimited  
fields terminated by ',';
```

```
insert into `customers` values (103,'Atelier graphique','Schmitt','Carine','40.32.2555','54, rue  
Royale',NULL,'Nantes',NULL,'44000','France',1370,'21000.00'),(112,'Signal Gift Stores','King','Jean','7025551838','8489 Strong  
St.',NULL,'Las Vegas','NV','83030','USA',1166,'71800.00'),(114,'Australian Collectors, Co.','Ferguson','Peter','03 9520 4555','636 St  
Kilda Road','Level 3','Melbourne','Victoria','3004','Australia',1611,'117300.00'),(119,'La Rochelle Gifts','Labruno','Janine  
,','40.67.8555','67, rue des Cinquante Otages',NULL,'Nantes',NULL,'44000','France',1370,'118200.00'),(121,'Baane Mini  
Imports','Bergulfsen','Jonas','07-98 9555','Erling Skakkas gate 78',NULL,'Stavern',NULL,'4110','Norway',1504,'81700.00'),(124,'Mini  
Gifts Distributors Ltd.','Nelson','Susan','4155551450','5677 Strong St.',NULL,'San  
Rafael','CA','97562','USA',1165,'210500.00'),(125,'Havel & Zbyszek Co.','Piestrzeniewicz','Zbyszek','(26) 642-7555','ul. Filtrowa  
68',NULL,'Warszawa',NULL,'01-012','Poland',NULL,'0.00'),(128,'Blauer See Auto, Co.','Keitel','Roland','+49 69 66 90 2555','Lyonerstr.  
34',NULL,'Frankfurt',NULL,'60528','Germany',1504,'59700.00'),(129,'Mini Wheels Co.','Murphy','Julie','6505555787','5557 North  
Pendale Street',NULL,'San Francisco','CA','94217','USA',1165,'64600.00'),(131,'Land of Toys Inc.','Lee','Kwai','2125557818','897 Long  
Airport Avenue',NULL,'NYC','NY','10022','USA',1323,'114900.00'),(141,'Euro+ Shopping Channel','Freyre','Diego','(91) 555 94 44','C/  
Moralzarzal, 86',NULL,'Madrid',NULL,'28034','Spain',1370,'227600.00'),(144,'Volvo Model Replicas, Co.','Berglund','Christina  
,','0921-12 3555','Berguvsvägen 8',NULL,'Luleå',NULL,'S-958 22','Sweden',1504,'53100.00'),(145,'Danish Wholesale  
Imports','Petersen','Jytte','31 12 3555','Vinbæltet 34',NULL,'Kobenhavn',NULL,'1734','Denmark',1401,'83400.00'),(146,'Saveley &  
Henriot, Co.','Saveley','Mary','78.32.5555','2, rue du Commerce',NULL,'Lyon',NULL,'69004','France',1337,'123900.00'),(148,'Dragon  
Souvenirs, Ltd.','Natividad','Eric','+65 221 7555','Bronz Sok','Bronz Apt. 3/6  
Teskivike','Singapore',NULL,'079903','Singapore',1621,'103800.00'),(151,'Muscle Machine Inc','Young','Jeff','2125557413','4092 Furth  
Circle','Suite 400','NYC','NY','10022','USA',1286,'138500.00'),(157,'Diecast Classics Inc.','Leong','Kelvin','2155551555','7586 Pompton  
St.',NULL,'Allentown','PA','70267','USA',1216,'100600.00'),(161,'Technics Stores Inc.','Hashimoto','Juri','6505556809','9408 Furth  
Circle',NULL,'Burlingame','CA','94217','USA',1165,'84600.00'),(166,'Handji Gifts & Co','Victorino','Wendy','+65 224 1555','106 Linden  
Road Sandown','2nd Floor','Singapore',NULL,'069045','Singapore',1612,'97900.00'),(167,'Herkku Gifts','Oeztan','Veyssel','+47 2267  
3215','Brehmen St. 121','PR 334 Sentrum','Bergen',NULL,'N 5804','Norway',1504,'96800.00'),(168,'American Souvenirs  
Inc','Franco','Keith','2035557845','149 Spinnaker Dr.','Suite 101','New Haven','CT','97823','USA',1286,'0.00'),(169,'Porto Imports  
Co.','de Castro','Isabel','(1) 356-5555','Estrada da saúde n. 58',NULL,'Lisboa',NULL,'1756','Portugal',NULL,'0.00'),(171,'Daedalus  
Designs Imports','Rancé','Martine','20.16.1555','184, chaussée de  
Tournai',NULL,'Lille',NULL,'59000','France',1370,'82900.00'),(172,'La Corne D\'abondance, Co.','Bertrand','Marie','(1)  
42.34.2555','265, boulevard Charonne',NULL,'Paris',NULL,'75012','France',1337,'84300.00'),(173,'Cambridge Collectables  
Co.','Tseng','Jerry','6175555555','4658 Baden Av.',NULL,'Cambridge','MA','51247','USA',1188,'43400.00'),(175,'Gift Depot  
Inc.','King','Julie','2035552570','25593 South Bay Ln.',NULL,'Bridgewater','CT','97562','USA',1323,'84300.00'),(177,'Osaka Souvenirs  
Co.','Kentary','Mory','+81 06 6342 5555','1-6-20 Dojima',NULL,'Kita-ku','Osaka','530-0003','Japan',1621,'81200.00'),(181,'Vitachrome  
Inc.','Frick','Michael','2125551500','2678 Kingston Rd.','Suite 101','NYC','NY','10022','USA',1286,'76400.00'),(186,'Toys of Finland,  
Co.','Karttunen','Matti','90-224 8555','Keskuskatu 45',NULL,'Helsinki',NULL,'21240','Finland',1501,'96500.00'),(187,'AV Stores,  
Co.','Ashworth','Rachel','(171) 555-1555','Fauntleroy Circus',NULL,'Manchester',NULL,'EC2 5NT','UK',1501,'136800.00'),(189,'Clover  
Collections, Co.','Cassidy','Dean','+353 1862 1555','25 Maiden Lane','Floor No.  
4','Dublin',NULL,'2','Ireland',1504,'69400.00'),(198,'Auto-Moto Classics Inc.','Taylor','Leslie','6175558428','16780 Pompton
```

St.',NULL,'Brickhaven','MA','58339','USA',1216,'23000.00'),(201,'UK Collectables, Ltd.','Devon','Elizabeth','(171) 555-2282','12, Berkeley Gardens Blvd',NULL,'Liverpool',NULL,'WX1 6LT','UK',1501,'92700.00'),(202,'Canadian Gift Exchange Network','Tamuri','Yoshi','(604) 555-3392','1900 Oak St.',NULL,'Vancouver','BC','V3F 2K1','Canada',1323,'90300.00'),(204,'Online Mini Collectables','Barajas','Miguel','6175557555','7635 Spinnaker Dr.',NULL,'Brickhaven','MA','58339','USA',1188,'68700.00'),(205,'Toys4GrownUps.com','Young','Julie','6265557265','78934 Hillside Dr.',NULL,'Pasadena','CA','90003','USA',1166,'90700.00'),(206,'Asian Shopping Network, Co','Walker','Brydey','+612 9411 1555','Suntec Tower Three','8 Temasek','Singapore',NULL,'038988','Singapore',NULL,'0.00'),(209,'Mini Caravy','Citeaux','Frédérique','88.60.1555','24, place Kléber',NULL,'Strasbourg',NULL,'67000','France',1370,'53800.00'),(211,'King Kong Collectables, Co.','Gao','Mike','+852 2251 1555','Bank of China Tower','1 Garden Road','Central Hong Kong',NULL,NULL,'Hong Kong',1621,'58600.00'),(216,'Enaco Distributors','Saavedra','Eduardo','(93) 203 4555','Rambla de Cataluña',23,NULL,'Barcelona',NULL,'08022','Spain',1702,'60300.00'),(219,'Boards & Toys Co.','Young','Mary','3105552373','4097 Douglas Av.',NULL,'Glendale','CA','92561','USA',1166,'11000.00'),(223,'Natürlich Autos','Kloss','Horst','0372-555188','Taucherstraße 10',NULL,'Cunewalde',NULL,'01307','Germany',NULL,'0.00'),(227,'Heintze Collectables','Ibsen','Palle','86 21 3555','Smagsloget 45',NULL,'Århus',NULL,'8200','Denmark',1401,'120800.00'),(233,'Québec Home Shopping Network','Fresnière','Jean','(514) 555-8054','43 rue St. Laurent',NULL,'Montréal','Québec','H3J 1C3','Canada',1286,'48700.00'),(237,'ANG Resellers','Camino','Alejandra','(91) 745 6555','Gran Vía, 1',NULL,'Madrid',NULL,'28001','Spain',NULL,'0.00'),(239,'Collectable Mini Designs Co.','Thompson','Valarie','7605558146','361 Furth Circle',NULL,'San Diego','CA','91217','USA',1166,'105000.00'),(240,'giftsbyemail.co.uk','Bennett','Helen','(198) 555-8888','Garden House','Crowther Way 23','Cowes','Isle of Wight','PO31 7PJ','UK',1501,'93900.00'),(242,'Alpha Cognac','Roulet','Annette','61.77.6555','1 rue Alsace-Lorraine',NULL,'Toulouse',NULL,'31000','France',1370,'61100.00'),(247,'Messner Shopping Network','Messner','Renate','069-0555984','Magazinweg 7',NULL,'Frankfurt',NULL,'60528','Germany',NULL,'0.00'),(249,'Amica Models & Co.','Accorti','Paolo','011-4988555','Via Monte Bianco 34',NULL,'Torino',NULL,'10100','Italy',1401,'113000.00'),(250,'Lyon Souvenirs','Da Silva','Daniel','+33 1 46 62 7555','27 rue du Colonel Pierre Avia',NULL,'Paris',NULL,'75508','France',1337,'68100.00'),(256,'Auto Associés & Cie.','Tonini','Daniel','30.59.8555','67, avenue de l'Europe',NULL,'Versailles',NULL,'78000','France',1370,'77900.00'),(259,'Toms Spezialitäten, Ltd','Pfalzheim','Henriette','0221-5554327','Mehrerheimerstr. 369',NULL,'Köln',NULL,'50739','Germany',1504,'120400.00'),(260,'Royal Canadian Collectables, Ltd.','Lincoln','Elizabeth','(604) 555-4555','23 Tsawassen Blvd.',NULL,'Tsawassen','BC','T2F 8M4','Canada',1323,'89600.00'),(273,'Franken Gifts, Co','Franken','Peter','089-0877555','Berliner Platz 43',NULL,'München',NULL,'80805','Germany',NULL,'0.00'),(276,'Anna's Decorations, Ltd','O'Hara','Anna','02 9936 8555','201 Miller Street','Level 15','North Sydney','NSW','2060','Australia',1611,'107800.00'),(278,'Rovelli Gifts','Rovelli','Giovanni','035-640555','Via Ludovico il Moro 22',NULL,'Bergamo',NULL,'24100','Italy',1401,'119600.00'),(282,'Souvenirs And Things Co.','Huxley','Adrian','+61 2 9495 8555','Monitor Money Building','815 Pacific Hwy','Chatswood','NSW','2067','Australia',1611,'93300.00'),(286,'Marta's Replicas Co.','Hernandez','Marta','6175558555','39323 Spinnaker Dr.',NULL,'Cambridge','MA','51247','USA',1216,'123700.00'),(293,'BG&E Collectables','Harrison','Ed','+41 26 425 50 01','Rte des Arsenaux 41',NULL,'Fribourg',NULL,'1700','Switzerland',NULL,'0.00'),(298,'Vida Sport, Ltd','Holz','Mihael','0897-034555','Grenzacherweg 237',NULL,'Genève',NULL,'1203','Switzerland',1702,'141300.00'),(299,'Norway Gifts By Mail, Co.','Klaeboe','Jan','+47 2212 1555','Drammensveien 126A','PB 211 Sentrum','Oslo',NULL,'N 0106','Norway',1504,'95100.00'),(303,'Schuyler Imports','Schuyler','Bradley','+31 20 491 9555','Kingsfordweg 151',NULL,'Amsterdam',NULL,'1043 GR','Netherlands',NULL,'0.00'),(307,'Der Hund Imports','Andersen','Mel','030-0074555','Obere Str. 57',NULL,'Berlin',NULL,'12209','Germany',NULL,'0.00'),(311,'Oulu Toy Supplies, Inc.','Koskitalo','Pirkko','981-443655','Torikatu 38',NULL,'Oulu',NULL,'90110','Finland',1501,'90500.00'),(314,'Petit Auto','Dewey','Catherine','(02) 5554 67','Rue Joseph-Bens 532',NULL,'Bruxelles',NULL,'B-1180','Belgium',1401,'79900.00'),(319,'Mini Classics','Frick','Steve','9145554562','3758 North Pendule Street',NULL,'White Plains','NY','24067','USA',1323,'102700.00'),(320,'Mini Creations Ltd.','Huang','Wing','5085559555','4575 Hillside Dr.',NULL,'New Bedford','MA','50553','USA',1188,'94500.00'),(321,'Corporate Gift Ideas Co.','Brown','Julie','6505551386','7734 Strong St.',NULL,'San Francisco','CA','94217','USA',1165,'105000.00'),(323,'Down Under Souvenirs, Inc','Graham','Mike','+64 9 312 5555','162-164 Grafton Road','Level 2','Auckland',NULL,NULL,'New Zealand',1612,'88000.00'),(324,'Stylish Desk Decors, Co.','Brown','Ann','(171) 555-0297','35 King George',NULL,'London',NULL,'WX3 6FW','UK',1501,'77000.00'),(328,'Tekni Collectables Inc.','Brown','William','2016559350','7476 Moss Rd.',NULL,'Newark','NJ','94019','USA',1323,'43000.00'),(333,'Australian Gift Network, Co','Calaghan','Ben','61-7-3844-6555','31 Duncan St. West End',NULL,'South Brisbane','Queensland','4101','Australia',1611,'51600.00'),(334,'Suominen Souvenirs','Suominen','Kalle','+358 9 8045 555','Software Engineering Center','SEC Oy','Espoo',NULL,'FIN-02271','Finland',1501,'98800.00'),(335,'Cramer Spezialitäten, Ltd','Cramer','Philip','0555-09555','Maubelstr. 90',NULL,'Brandenburg',NULL,'14776','Germany',NULL,'0.00'),(339,'Classic Gift Ideas, Inc','Cervantes','Francisca','2155554695','782 First Street',NULL,'Philadelphia','PA','71270','USA',1188,'81100.00'),(344,'CAF Imports','Fernandez','Jesus','+34 913 728 555','Merchants House','27-30 Merchant's Quay','Madrid',NULL,'28023','Spain',1702,'59600.00'),(347,'Men \R\ US Retailers, Ltd.','Chandler','Brian','2155554369','6047 Douglas Av.',NULL,'Los Angeles','CA','91003','USA',1166,'57700.00'),(348,'Asian Treasures, Inc.','McKenna','Patricia','2967 555','8 Johnstown Road',NULL,'Cork','Co. Cork',NULL,'Ireland',NULL,'0.00'),(350,'Marseille Mini Autos','Lebihan','Laurence','91.24.4555','12, rue des Bouchers',NULL,'Marseille',NULL,'13008','France',1337,'65000.00'),(353,'Reims Collectables','Henriot','Paul','26.47.1555','59 rue de l'Abbaye',NULL,'Reims',NULL,'51100','France',1337,'81100.00'),(356,'SAR Distributors, Co','Kuger','Armand','+27 21 550 3555','1250 Pretorius Street',NULL,'Hatfield','Pretoria','0028','South Africa',NULL,'0.00'),(357,'GiftsForHim.com','MacKinlay','Wales','64-9-3763555','199 Great North Road',NULL,'Auckland',NULL,NULL,'New Zealand',1612,'77700.00'),(361,'Kommission Auto','Josephs','Karin','0251-555259','Luisenstr. 48',NULL,'Münster',NULL,'44087','Germany',NULL,'0.00'),(362,'Gifts4AllAges.com','Yoshido','Juri','6175559555','8616 Spinnaker Dr.',NULL,'Boston','MA','51003','USA',1216,'41900.00'),(363,'Online Diecast Creations Co.','Young','Dorothy','6035558647','2304 Long Airport Avenue',NULL,'Nashua','NH','62016','USA',1216,'114200.00'),(369,'Lisboa Souvenirs, Inc','Rodriguez','Lino','(1) 354-2555','Jardim das rosas n. 32',NULL,'Lisboa',NULL,'1675','Portugal',NULL,'0.00'),(376,'Precious

Collectables','Urs','Braun','0452-076555','Hauptstr. 29',NULL,'Bern',NULL,'3012','Switzerland',1702,'0.00'),(379,'Collectables For Less Inc.','Nelson','Allen','6175558555','7825 Douglas Av.',NULL,'Brickhaven','MA','58339','USA',1188,'70700.00'),(381,'Royale Belge','Cartrain','Pascal','(071) 23 67 2555','Boulevard Tirou, 255',NULL,'Charleroi',NULL,'B-6000','Belgium',1401,'23500.00'),(382,'Salzburg Collectables','Pipps','Georg','6562-9555','Geislweg 14',NULL,'Salzburg',NULL,'5020','Austria',1401,'71700.00'),(385,'Cruz & Sons Co.','Cruz','Arnold','+63 2 555 3587','15 McCallum Street','NatWest Center #13-03','Makati City',NULL,'1227 MM','Philippines',1621,'81500.00'),(386,'L'ordine Souvenirs','Moroni','Maurizio','0522-556555','Strada Provinciale 124',NULL,'Reggio Emilia',NULL,'42100','Italy',1401,'121400.00'),(398,'Tokyo Collectables Ltd','Shimamura','Akiko','+81 3 3584 0555','2-2-8 Roppongi',NULL,'Minato-ku','Tokyo','106-0032','Japan',1621,'94400.00'),(406,'Auto Canal+ Petit','Perrier','Dominique','(1) 47.55.6555','25, rue Lauriston',NULL,'Paris',NULL,'75016','France',1337,'95000.00'),(409,'Stuttgart Collectable Exchange','Müller','Rita','0711-555361','Adenauerallee 900',NULL,'Stuttgart',NULL,'70563','Germany',NULL,'0.00'),(412,'Extreme Desk Decorations, Ltd','McRoy','Sarah','04 499 9555','101 Lambton Quay','Level 11','Wellington',NULL,NULL,'New Zealand',1612,'86800.00'),(415,'Bavarian Collectables Imports, Co.','Donnermeyer','Michael','+49 89 61 08 9555','Hansastr. 15',NULL,'Munich',NULL,'80686','Germany',1504,'77000.00'),(424,'Classic Legends Inc.','Hernandez','Maria','2125558493','5905 Pompton St.','Suite 750','NYC','NY','10022','USA',1286,'67500.00'),(443,'Feuer Online Stores, Inc.','Feuer','Alexander','0342-555176','Heerstr. 22',NULL,'Leipzig',NULL,'04179','Germany',NULL,'0.00'),(447,'Gift Ideas Corp.','Lewis','Dan','2035554407','2440 Pompton St.',NULL,'Glendale','CT','97561','USA',1323,'49700.00'),(448,'Scandinavian Gift Ideas','Larsson','Martha','0695-34 6555','Åkergatan 24',NULL,'Bräcke',NULL,'S-844 67','Sweden',1504,'116400.00'),(450,'The Sharp Gifts Warehouse','Frick','Sue','4085553659','3086 Ingle Ln.',NULL,'San Jose','CA','94217','USA',1165,'77600.00'),(452,'Mini Auto Werke','Mendel','Roland','7675-3555','Kirchgasse 6',NULL,'Graz',NULL,'8010','Austria',1401,'45300.00'),(455,'Super Scale Inc.','Murphy','Leslie','2035559545','567 North Pendale Street',NULL,'New Haven','CT','97823','USA',1286,'95400.00'),(456,'Microscale Inc.','Choi','Yu','2125551957','5290 North Pendale Street','Suite 200','NYC','NY','10022','USA',1286,'39800.00'),(458,'Corrida Auto Replicas, Ltd','Sommer','Martin','(91) 555 22 82','C/ Araquil, 67',NULL,'Madrid',NULL,'28023','Spain',1702,'104600.00'),(459,'Warburg Exchange','Ottlieb','Sven','0241-039123','Walsersweg 21',NULL,'Aachen',NULL,'52066','Germany',NULL,'0.00'),(462,'FunGiftIdeas.com','Benitez','Violeta','5085552555','1785 First Street',NULL,'New Bedford','MA','50553','USA',1216,'85800.00'),(465,'Anton Designs, Ltd.','Anton','Carmen','+34 913 728555','c/ Gobelas, 19-1 Urb. La Florida',NULL,'Madrid',NULL,'28023','Spain',NULL,'0.00'),(471,'Australian Collectables, Ltd','Clenahan','Sean','61-9-3844-6555','7 Allen Street',NULL,'Glen Waverly','Victoria','3150','Australia',1611,'60300.00'),(473,'Frau da Collezione','Ricotti','Franco','+39 022515555','20093 Cologno Monzese','Alessandro Volta 16','Milan',NULL,NULL,'Italy',1401,'34800.00'),(475,'West Coast Collectables Co.','Thompson','Steve','3105553722','3675 Furth Circle',NULL,'Burbank','CA','94019','USA',1166,'55400.00'),(477,'Mit Vergnügen & Co.','Moos','Hanna','0621-08555','Forsterstr. 57',NULL,'Mannheim',NULL,'68306','Germany',NULL,'0.00'),(480,'Kremlin Collectables, Co.','Semenov','Alexander','+7 812 293 0521','2 Pobedy Square',NULL,'Saint Petersburg',NULL,'196143','Russia',NULL,'0.00'),(481,'Raanan Stores, Inc','Altagar,G M','Raanan','+ 972 9 959 8555','3 Hagalim Blv.',NULL,'Herzlia',NULL,'47625','Israel',NULL,'0.00'),(484,'Iberia Gift Imports, Corp.','Roel','José Pedro','(95) 555 82 82','C/ Romero, 33',NULL,'Sevilla',NULL,'41101','Spain',1702,'65700.00'),(486,'Motor Mint Distributors Inc.','Salazar','Rosa','2155559857','11328 Douglas Av.',NULL,'Philadelphia','PA','71270','USA',1323,'72600.00'),(487,'Signal Collectibles Ltd.','Taylor','Sue','4155554312','2793 Furth Circle',NULL,'Brisbane','CA','94217','USA',1165,'60300.00'),(489,'Double Decker Gift Stores, Ltd','Smith','Thomas','(171) 555-7555','120 Hanover Sq.',NULL,'London',NULL,'WA1 1DP','UK','1501','43300.00'),(495,'Diecast Collectables','Franco','Valarie','6175552555','6251 Ingle Ln.',NULL,'Boston','MA','51003','USA',1188,'85100.00'),(496,'Kelly's Gift Shop','Snowden','Tony','+64 9 5555500','Arenales 1938 3\A\','NULL','Auckland',NULL,NULL,'New Zealand',1612,'110000.00');

CREATE TABLE `payments` (`customerNumber` int , `checkNumber` varchar(50) ,
`paymentDate` date , `amount` decimal(10,2)) row format delimited fields terminated
by ';;';

insert into `payments` values

(103,'HQ336336','2016-10-19','6066.78'),(103,'JM555205','2016-10-05','14571.44'),(103,'OM314933','2016-10-18','1676.14'),(112,'B O864823','2016-10-17','14191.12'),(112,'HQ55022','2016-10-06','32641.98'),(112,'ND748579','2016-10-20','33347.88'),(114,'GG3145 5','2016-10-20','45864.03'),(114,'MA765515','2016-10-15','82261.22'),(114,'NP603840','2016-10-31','7565.08'),(114,'NR27552','2016 -10-10','44894.74'),(119,'DB933704','2016-10-14','19501.82'),(119,'LN373447','2016-10-08','47924.19'),(119,'NG94694','2016-10-22','49523.67'),(121,'DB889831','2016-10-16','50218.95'),(121,'FD317790','2016-10-28','1491.38'),(121,'KI831359','2016-10-04','17876.3 2'),(121,'MA302151','2016-10-28','34638.14'),(124,'AE215433','2016-10-05','101244.59'),(124,'BG255406','2016-10-28','85410.87'),(124,'CQ287967','2016-10-11','11044.30'),(124,'ET64396','2016-10 -16','83598.04'),(124,'HI366474','2016-10-27','47142.70'),(124,'HR86578','2016-10-02','55639.66'),(124,'KI131716','2016-10-15','111 654.40'),(124,'LF217299','2016-10-26','43369.30'),(124,'NT141748','2016-10-25','45084.38'),(128,'DI925118','2016-10-28','10549.01 '), (128,'FA465482','2016-10-18','24101.81'),(128,'FH668230','2016-10-24','33820.62'),(128,'IP383901','2016-10-18','7466.32'),(129,'D M826140','2016-10-08','26248.78'),(129,'ID449593','2016-10-11','23923.93'),(129,'PI42991','2016-10-09','16537.85'),(131,'CL442705 ','2016-10-12','22292.62'),(131,'MA724562','2016-10-02','50025.35'),(131,'NB445135','2016-10-11','35321.97'),(141,'AU364101','201 6-10-19','36251.03'),(141,'DB583216','2016-10-01','36140.38'),(141,'DL460618','2016-10-19','46895.48'),(141,'HJ32686','2016-10-30','59830.55'),(141,'ID10962','2016-10-31','116208.40'),(141,'IN446258','2016-10-25','65071.26'),(141,'JE105477','2016-10-18','120166 .58'),(141,'JN355280','2016-10-26','49539.37'),(141,'JN722010','2016-10-25','40206.20'),(141,'KT52578','2016-10-09','63843.55'),(14 1,'MC46946','2016-10-09','35420.74'),(141,'MF629602','2016-10-16','20009.53'),(141,'NU627706','2016-10-17','26155.91'),(144,'IR8 46303','2016-10-12','36005.71'),(144,'LA685678','2016-10-09','7674.94'),(145,'CN328545','2016-10-03','4710.73'),(145,'ED39322','20 16-10-26','28211.70'),(145,'HR182688','2016-10-01','20564.86'),(145,'JJ246391','2016-10-20','53959.21'),(146,'FP549817','2016-10-1

8','40978.53'),(146,'FU793410','2016-10-16','49614.72'),(146,'LJ160635','2016-10-10','39712.10'),(148,'BI507030','2016-10-22','4438 0.15'),(148,'DD635282','2016-10-11','2611.84'),(148,'KM172879','2016-10-26','105743.00'),(148,'ME497970','2016-10-27','3516.04'),(151,'BF686658','2016-10-22','58793.53'),(151,'GB852215','2016-10-26','20314.44'),(151,'IP568906','2016-10-18','58841.35'),(151,'KI 884577','2016-10-14','39964.63'),(157,'H1618861','2016-10-19','35152.12'),(157,'NN711988','2016-10-07','63357.13'),(161,'BR35238 4','2016-10-14','2434.25'),(161,'BR478494','2016-10-18','50743.65'),(161,'KG644125','2016-10-02','12692.19'),(161,'NI908214','2016- 10-05','38675.13'),(166,'BQ327613','2016-10-16','38785.48'),(166,'DC979307','2016-10-07','44160.92'),(166,'LA318629','2016-10-28',' 22474.17'),(167,'ED743615','2016-10-19','12538.01'),(167,'GN228846','2016-10-03','85024.46'),(171,'GB878038','2016-10-15','1899 7.89'),(171,'IL104425','2016-10-22','42783.81'),(172,'AD832091','2016-10-09','1960.80'),(172,'CE51751','2016-10-04','51209.58'),(17 2,'EH208589','2016-10-20','33383.14'),(173,'GP545698','2016-10-13','11843.45'),(173,'IG462397','2016-10-29','20355.24'),(175,'CITI3 434344','2016-10-19','28500.78'),(175,'IO448913','2016-10-19','24879.08'),(175,'PI15215','2016-10-10','42044.77'),(177,'AU750837',' 2016-10-17','15183.63'),(177,'CI381435','2016-10-19','47177.59'),(181,'CM564612','2016-10-25','22602.36'),(181,'GQ132144','2016- 10-30','5494.78'),(181,'OH367219','2016-10-16','44400.50'),(186,'AE192287','2016-10-10','23602.90'),(186,'AK412714','2016-10-27',' 37602.48'),(186,'KA602407','2016-10-21','34341.08'),(187,'AM968797','2016-10-03','52825.29'),(187,'BQ39062','2016-10-08','47159. 11'),(187,'KL124726','2016-10-27','48425.69'),(189,'BO711618','2016-10-03','17359.53'),(189,'NM916675','2016-10-01','32538.74'),(198,'FI192930','2016-10-06','9658.74'),(198,'HQ920205','2016-10-06','6036.96'),(198,'IS946883','2016-10-21','5858.56'),(201,'DP67 013','2016-10-20','23908.24'),(201,'OO846801','2016-10-15','37258.94'),(202,'HI358554','2016-10-18','36527.61'),(202,'IQ627690','2 016-10-08','33594.58'),(204,'GC697638','2016-10-13','51152.86'),(204,'IS150005','2016-10-24','4424.40'),(205,'GL756480','2016-10-0 4','3879.96'),(205,'LL562733','2016-10-05','50342.74'),(205,'NM739638','2016-10-06','39580.60'),(209,'BOAF82044','2016-10-03','35 157.75'),(209,'ED520529','2016-10-21','4632.31'),(209,'PH785937','2016-10-04','36069.26'),(211,'BJ535230','2016-10-09','45480.79'), (216,'BG407567','2016-10-09','3101.40'),(216,'ML780814','2016-10-06','24945.21'),(216,'MM342086','2016-10-14','40473.86'),(219,' BN17870','2016-10-02','3452.75'),(219,'BR941480','2016-10-18','4465.85'),(227,'MQ413968','2016-10-31','36164.46'),(227,'NU21326 ', '2016-10-02','53745.34'),(233,'BOFA23232','2016-10-20','29070.38'),(233,'II180006','2016-10-01','22997.45'),(233,'JG981190','2016- 10-18','16909.84'),(239,'NQ865547','2016-10-15','80375.24'),(240,'IF245157','2016-10-16','46788.14'),(240,'JO719695','2016-10-28',' 24995.61'),(242,'AF40894','2016-10-22','33818.34'),(242,'HR224331','2016-10-03','12432.32'),(242,'KI744716','2016-10-21','14232.7 0'),(249,'IJ399820','2016-10-19','33924.24'),(249,'NE404084','2016-10-04','48298.99'),(250,'EQ12267','2016-10-17','17928.09'),(250,' HD284647','2016-10-30','26311.63'),(250,'HN114306','2016-10-18','23419.47'),(256,'EP227123','2016-10-10','5759.42'),(256,'HE8493 6','2016-10-22','53116.99'),(259,'EU280955','2016-10-06','61234.67'),(259,'GB361972','2016-10-07','27988.47'),(260,'IO164641','201 6-10-30','37527.58'),(260,'NH776924','2016-10-24','29284.42'),(276,'EM979878','2016-10-09','27083.78'),(276,'KM841847','2016-10- 13','38547.19'),(276,'LE432182','2016-10-28','41554.73'),(276,'OJ819725','2016-10-30','29848.52'),(278,'BJ483870','2016-10-05','376 54.09'),(278,'GP636783','2016-10-02','52151.81'),(278,'NI983021','2016-10-24','37723.79'),(282,'IA793562','2016-10-03','24013.52'), (282,'JT819493','2016-10-02','35806.73'),(282,'OD327378','2016-10-03','31835.36'),(286,'DR578578','2016-10-28','47411.33'),(286,'K H910279','2016-10-05','43134.04'),(298,'AJ574927','2016-10-13','47375.92'),(298,'LF501133','2016-10-18','61402.00'),(299,'AD30408 5','2016-10-24','36798.88'),(299,'NR157385','2016-10-05','32260.16'),(311,'DG336041','2016-10-15','46770.52'),(311,'FA728475','201 6-10-06','32723.04'),(311,'NQ966143','2016-10-25','16212.59'),(314,'LQ244073','2016-10-09','45352.47'),(314,'MD809704','2016-10- 03','16901.38'),(319,'HL685576','2016-10-06','42339.76'),(319,'OM548174','2016-10-07','36092.40'),(320,'GJ597719','2016-10-18','8 307.28'),(320,'HO576374','2016-10-20','41016.75'),(320,'MU817160','2016-10-24','52548.49'),(321,'DJ15149','2016-10-03','85559.12 '), (321,'LA556321','2016-10-15','46781.66'),(323,'AL493079','2016-10-23','75020.13'),(323,'ES347491','2016-10-24','37281.36'),(323,' HG738664','2016-10-05','2880.00'),(323,'PQ803830','2016-10-24','39440.59'),(324,'DQ409197','2016-10-13','13671.82'),(324,'FP443 161','2016-10-07','29429.14'),(324,'HB150714','2016-10-23','37455.77'),(328,'EN930356','2016-10-16','7178.66'),(328,'NR631421','2 016-10-30','31102.85'),(333,'HL209210','2016-10-15','23936.53'),(333,'JK479662','2016-10-17','9821.32'),(333,'NF959653','2016-10-0 1','21432.31'),(334,'CS435306','2016-10-27','45785.34'),(334,'HH517378','2016-10-16','29716.86'),(334,'LF737277','2016-10-22','283 94.54'),(339,'AP286625','2016-10-24','23333.06'),(339,'DA98827','2016-10-28','34606.28'),(344,'AF246722','2016-10-24','31428.21'),(344,'NJ906924','2016-10-02','15322.93'),(347,'DG700707','2016-10-18','21053.69'),(347,'LG808674','2016-10-24','20452.50'),(350,'B Q602907','2016-10-11','18888.31'),(350,'CI471510','2016-10-25','50824.66'),(350,'OB648482','2016-10-29','1834.56'),(353,'CO35119 3','2016-10-10','49705.52'),(353,'ED878227','2016-10-21','13920.26'),(353,'GT878649','2016-10-21','16700.47'),(353,'HU618252','201 6-10-09','46656.94'),(357,'AG240323','2016-10-16','20220.04'),(357,'NB291497','2016-10-15','36442.34'),(362,'FP170292','2016-10-1 1','18473.71'),(362,'OG208861','2016-10-21','15059.76'),(363,'HL575273','2016-10-17','50799.69'),(363,'IS232033','2016-10-16','102 23.83'),(363,'PN238558','2016-10-05','55425.77'),(379,'CA762595','2016-10-12','28322.83'),(379,'FR499138','2016-10-16','32680.31'), (379,'GB890854','2016-10-02','12530.51'),(381,'BC726082','2016-10-03','12081.52'),(381,'CC475233','2016-10-19','1627.56'),(381,'G B117430','2016-10-03','14379.90'),(381,'MS154481','2016-10-22','1128.20'),(382,'CC871084','2016-10-12','35826.33'),(382,'CT82114 7','2016-10-01','6419.84'),(382,'PH29054','2016-10-27','42813.83'),(385,'BN347084','2016-10-02','20644.24'),(385,'CP804873','2016- 10-19','15822.84'),(385,'EK785462','2016-10-09','51001.22'),(386,'DO106109','2016-10-18','38524.29'),(386,'HG438769','2016-10-18', '51619.02'),(398,'AJ478695','2016-10-14','33967.73'),(398,'DO787644','2016-10-21','22037.91'),(398,'JPMR4544','2016-10-18','615.4 5'),(398,'KB54275','2016-10-29','48927.64'),(406,'BJMPR4545','2016-10-23','12190.85'),(406,'HJ217687','2016-10-28','49165.16'),(40 6,'NA197101','2016-10-17','25080.96'),(412,'GH197075','2016-10-25','35034.57'),(412,'PJ434867','2016-10-14','31670.37'),(415,'ER5 4537','2016-10-28','31310.09'),(424,'KF480160','2016-10-07','25505.98'),(424,'LM271923','2016-10-16','21665.98'),(424,'OA595449', '2016-10-31','22042.37'),(447,'AO757239','2016-10-15','6631.36'),(447,'ER615123','2016-10-25','17032.29'),(447,'OU516561','2016-1 0-17','26304.13'),(448,'FS299615','2016-10-18','27966.54'),(448,'KR822727','2016-10-30','48809.90'),(450,'EF485824','2016-10-21','5 9551.38'),(452,'ED473873','2016-10-15','27121.90'),(452,'FN640986','2016-10-20','15130.97'),(452,'HG635467','2016-10-03','8807.1 2'),(455,'HA777606','2016-10-05','38139.18'),(455,'IR662429','2016-10-12','32239.47'),(456,'GJ715659','2016-10-13','27550.51'),(456 ', 'MO743231','2016-10-30','1679.92'),(458,'DD995006','2016-10-15','33145.56'),(458,'NA377824','2016-10-06','22162.61'),(458,'OO6 06861','2016-10-13','57131.92'),(462,'ED203908','2016-10-15','30293.77'),(462,'GC60330','2016-10-08','9977.85'),(462,'PE176846','2 016-10-27','48355.87'),(471,'AB661578','2016-10-28','9415.13'),(471,'CO645196','2016-10-10','35505.63'),(473,'LL427009','2016-10- 17','7612.06'),(473,'PC688499','2016-10-27','17746.26'),(475,'JP113227','2016-10-09','7678.25'),(475,'PB951268','2016-10-13','36070 .47'),(484,'GK294076','2016-10-26','3474.66'),(484,'JH546765','2016-10-29','47513.19'),(486,'BL66528','2016-10-14','5899.38'),(486,' HS86661','2016-10-23','45994.07'),(486,'JB117768','2016-10-20','25833.14'),(487,'AH612904','2016-10-28','29997.09'),(487,'PT55018

```
1','2016-10-29','12573.28'),(489,'OC773849','2016-10-04','22275.73'),(489,'PO860906','2016-10-31','7310.42'),(495,'BH167026','2016-10-26','59265.14'),(495,'FN155234','2016-10-14','6276.60'),(496,'EU531600','2016-10-25','30253.75'),(496,'MB342426','2016-10-16','32077.44'),(496,'MN89921','2016-10-31','52166.01');
```

1. How to create an empty table in hive from another table without copying data?

```
use retail;
create table paymentscopy as select * from payments where 1=2;
or
show create table payments;
```

2. Can we join data between tables residing in two different databases?

Yes, prefix the schema before every table name.

3. Can we insert data into hive table? multiple records in one time without using load command?

Yes, refer the script mentioned above.

Duplicate handling:

(Lets insert 9 rows duplicate in payments)

```
insert into `payments` values
(103,'HQ336336','2016-10-19','6066.78'),(103,'JM555205','2016-10-05','14571.44'),(103,'OM314933','2016-10-18','1676.14'),(112,'BO864823','2016-10-17','14191.12'),(112,'HQ55022','2016-10-06','32641.98'),(114,'GG31455','2016-10-20','45864.03'),(114,'MA765515','2016-10-15','82261.22'),(114,'NP603840','2016-10-31','7565.08'),(114,'NR27552','2016-10-10','44894.74');
```

(Lets insert 3 rows duplicate into customers)

```
insert into `customers` values (103,'Atelier graphique','Schmitt','Carine ','40.32.2555','54, rue Royale',NULL,'Nantes',NULL,'44000','France',1370,'21000.00'),(112,'Signal Gift Stores','King','Jean','7025551838','8489 Strong St.',NULL,'Las Vegas','NV','83030','USA',1166,'71800.00'),(114,'Australian Collectors, Co.','Ferguson','Peter','03 9520 4555','636 St Kilda Road','Level 3','Melbourne','Victoria','3004','Australia',1611,'117300.00'),(114,'Australian Collectors, Co.','Ferguson','Peter','03 9520 4555','636 St Kilda Road','Level 3','Melbourne','Victoria','3004','Australia',1611,'117300.00');
```

```
select * from payments;
```

```
select * from customers;
```

4. How to identify the duplicate records in a table?

When you use group by is the aggregation is mandatory? No

Payments -> we need to get 282 - 9 rows = 273

Customers -> we need to get 126-4 =122

Option 1:

```
select distinct * from payments; ---22.60 seconds for 274 records
```

Option 2:

```
select customerNumber,checkNumber,paymentDate,amount
from payments
```

group by customerNumber,checkNumber,paymentDate,amount; ---20.45 seconds for 274 records

5. **How to identify which are the customer table records are duplicated and how many duplicates are there?**
6. **When do you use where and when do you use having for filtering of data?
where used for direct filter, where having used for aggregated filter.**

```
select customernumber, customername, contactlastname, contactfirstname, phone, addressline1,
addressline2, city, state, postalcode, country, salesrepemployeenumber, creditlimit,count(1)
from customers
group by customernumber, customername, contactlastname, contactfirstname, phone, addressline1,
addressline2, city, state, postalcode, country, salesrepemployeenumber, creditlimit
having count(1)>1;
```

7. **How to remove duplicate payments and retain only the de duplicated payment information in a table with or without partition?
How do you delete duplicate data from a hive table?**

Regular Table with less volume of data:

```
--don't run this insert - insert overwrite table payments select distinct * from payments;
select * from payments_dup;
```

```
create table payments_dup as select * from payments;
select * from payments_dup;
```

```
insert overwrite table payments_dup select distinct * from payments_dup;
select * from payments_dup;
```

Partitioned Table in a big table:

```
CREATE TABLE `payments_part` ( `customerNumber` int , `checkNumber` varchar(50) , `amount`
decimal(10,2) ) partitioned by (`paymentDate` date) row format delimited fields terminated by ',';
```

```
set hive.exec.dynamic.partition.mode=nonstrict
```

```
insert overwrite table payments_part partition (paymentdate) select
customernumber,checknumber,amount,paymentdate from payments;
```

```
select customernumber,checknumber,amount,paymentdate,count(1) from payments_part group by
customernumber,checknumber,amount,paymentdate having count(1)>1;
```

```
insert overwrite table payments_part partition (paymentdate)
select distinct customernumber,checknumber,amount,paymentdate
from payments_part;
```

```
select customernumber,checknumber,amount,paymentdate,count(1)
from payments_part
group by customernumber,checknumber,amount,paymentdate
having count(1)>1;
```

8. Show the customer info who made first payment to our company?

```
select * from payments_part where paymentdate in (select min(paymentdate) from payments_part );
```

9. Show the first payment, last made by the customer?

```
select * from payments_part where customernumber=496 and paymentdate in (select min(paymentdate) from payments_part where customernumber=496);
```

```
select a.* from payments_part as a where a.customernumber=496 and a.paymentdate in (select max(b.paymentdate) from payments_part as b where b.customernumber=496);
```

10. Show the last (but one) or second recent payment made by the customer 496?

This is the way u can achieve the result, but not a right way to do, in hive the nested subquery with more than 1 is not supported.

Nested SubQuery expressions are not supported.

Multicolumn sub query also not supported.

```
select a.* from payments_part as a
where a.customernumber=496 and a.paymentdate in (select max(b.paymentdate) from payments_part as b
where b.customernumber=496 and b.paymentdate not in (select max(c.paymentdate) from
payments_part as c where c.customernumber=496));
```

```
FAILED: SemanticException [Error 10249]: Line 2:151
Unsupported SubQuery Expression 'paymentdate': Nested SubQuery
expressions are not supported.
```

Windowing/Analytical Functions:

11. Show the lowest payment, highest, number of payment made by the customer 496?

```
insert into `payments_part` partition(paymentdate='2016-10-30') values (496,'HQ336436',52166.01);
```

Cume_dist

It returns the cumulative distribution of a value. It results from 0 to 1. For suppose if the total number of records are 10 then for the 1st row the cume_dist will be 1/10 and for the second 2/10 and so on till 10/10.

Rank

The rank function will return the rank of the values as per the result set of the over clause. If two values are same then it will give the same rank to those 2 values and then for the next value, the sub-sequent rank will be skipped.

Row_number

Row number will return the continuous sequence of numbers for all the rows of the result set of the over clause.

Dense_rank

It is same as the rank() function but the difference is if any duplicate value is present then the rank will not be skipped for the subsequent rows. Each unique value will get the ranks in a sequence.

```
set hive.cli.print.header=true;
select customernumber,paymentdate,amount, rank() over(partition by customernumber order by amount
desc) as rnk,dense_rank() over(partition by customernumber order by amount desc) as
d_rank,row_number() over(partition by customernumber order by amount desc) as rownum,cume_dist()
over(partition by customernumber order by paymentdate) as cumulative_dist from payments_part where
customernumber=496;
```

EMPID	Name	Salary	_Rank	DenseRank	RowNumber
EMP102	Sam	20000.00	1	1	1
EMP108	Rahul	20000.00	1	1	2
EMP111	akshay	20000.00	1	1	3
EMP106	Maresh	18000.00	4	2	4
EMP101	Vishal	15000.00	5	3	5
EMP109	menaka	15000.00	5	3	6
EMP105	Ravi	10000.00	7	4	7

12. Show the second, third largest payment made by the customer 496?

```
select customernumber,paymentdate,amount,rownumdt from
(select customernumber,paymentdate,amount, rank() over(partition by customernumber order by
amount desc) as rnk,dense_rank() over(partition by customernumber order by amount desc) as
d_rank,row_number() over(partition by customernumber order by amount desc) as
rownum,cume_dist() over(partition by customernumber order by paymentdate) as
cumulative_dist,row_number() over(partition by customernumber order by paymentdate desc) as
rownumdt from payments_part where customernumber=496) as temp
where rownum in (2,3);
```

13. Show the customer purchase rate whether growing up or leaning down from the largest and smallest payment made?

```
select customernumber,paymentdate,amount,max_amt-amount as maxamt,case when min_amt-amount
<0 then "purchase capacity is improved" when min_amt-amount >0 then "purchase capacity is reduced"
else "" end as minamt from (select customernumber,paymentdate,amount, max(amount) over(partition
by customernumber) as max_amt,min(amount) over(partition by customernumber) as min_amt
from payments_part
where customernumber=496) as temp;
```

14. Show the customer purchase rate whether growing up or leaning down from the immediate previous and next payment made?

```
--case when condition then exec when cond2 then exe else exec end as alias
select customernumber,paymentdate,amount, case when amount_paid_previous_day>= amount then
"prior payment is high" else "prior payment is small" end as lag,
case when amount_paid_next_day>= amount then "next day payment is high" else "next day payment is
small" end as lead
from(select customernumber, paymentdate,lag(amount) over(partition by customernumber order by
paymentdate) as amount_paid_previous_day,amount, lead(amount) over(partition by customernumber
order by paymentdate) as amount_paid_next_day from payments_part where customernumber in
(103,496)) as temp;
```

15. How to Create random unique numbers?

```
select regexp_replace(reflect('java.util.UUID','randomUUID'),'-', '') as row_num,p.* from payments as p;
```

16. How to create surrogate key or sequence number?


```
insert into table payments_increment select row_number() over() +
coalesce(maxseq,0),customernumber,checknumber,paymentdate,amount
from payments as p
left outer join (select 1 as dummy,max(seqnum) maxseq from payments_increment) as pi
on pi.dummy=1;
```

```
select * from payments_increment where customernumber=496;
```

17. How to create version number for the payments made by the customers?

current data

cid ver amt

10 1 1000

10 2 1100

10 3 3000

10 4 2000

10 5 3050

next day

cid amt

10 2000

10 3050

```
CREATE TABLE `payments_version` ( `version` int, `customerNumber` int, `checkNumber` varchar(50) ,
`paymentDate` date , `amount` decimal(10,2) ) row format delimited fields terminated by ',';
```

```
insert overwrite table payments_version select row_number() over(partition by customernumber order by
paymentdate),customernumber,checknumber,paymentdate,amount from payments;
```

```
insert into table payments_version select row_number() over(partition by p.customernumber order by
paymentdate) + coalesce(maxversion,0),p.customernumber,checknumber,paymentdate,amount
from payments as p
left outer join (select customernumber,max(version) maxversion from payments_version group by
customernumber) as pi
on pi.customernumber=p.customernumber;
```

```
select * from payments_version where customernumber=496;
```

Joins: Inner, Outer (left, right, full), semi, anti, self, cross

```
insert into `payments` values
```

```
(1000,'HQ336336','2016-10-19','6066.78'),(1003,'JM555205','2016-10-05','14571.44');
```

18. Show me only the customers who made the payments and what is the payment amount (this is possible only with join).

Join/subquery/exists can be used, but for multiple columns comparison for eg. customernumber and customerphonenumber we cant use subquery, we can only use join in the case if we need some columns from payments table also.

```
select c.customernumber,c.customername,p.amount from customers c inner join payments p on
c.customernumber=p.customernumber;
```

select c.customernumber,c.customername from customers c where exists(select p.customernumber from payments p where p.customernumber=c.customernumber);

select c.customernumber,c.customername from customers c where c.customernumber in (select customernumber from payments p);

19. Show me only the customers who didnt make the payments.

select c.customernumber,c.customername,p.amount from customers c **left** join payments p on c.customernumber=p.customernumber where p.customernumber is null;

select c.customernumber,c.customername from customers c where c.customernumber not in (select customernumber from payments p);

select c.customernumber,c.customername from customers c where not exists(select p.customernumber from payments p where p.customernumber=c.customernumber);

20. Show me the customers with no details who made the payments anonymously

select c.customernumber,c.customername,p.amount from customers c **right** outer join payments p on c.customernumber=p.customernumber where c.customernumber is null;

21. show the customers and payments info in both the above cases

select c.customernumber,c.customername,p.amount from customers c **full** outer join payments p on c.customernumber=p.customernumber;

22. Show me the customers only once who paid the payments

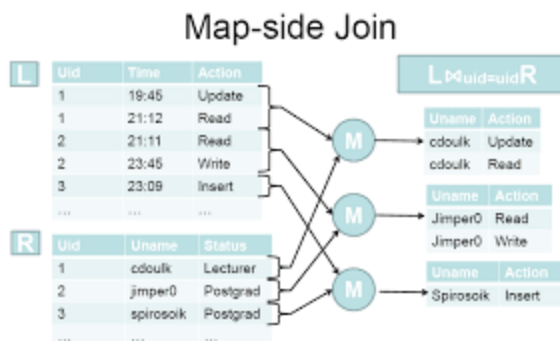
An INNER JOIN returns the columns from both tables. A LEFT SEMI JOIN only returns the records from the left-hand table, If there are multiple matching rows in the right-hand column, an INNER JOIN will return one row for each match on the right table, while a LEFT SEMI JOIN only returns the rows from the left table, regardless of the number of matching rows on the right side.

select c.customernumber from payments c **left semi join** customers p on (c.customernumber=p.customernumber);

select c.customernumber,c.customername,p.amount from customers c inner join (select max(customernumber)payments) as p on (c.customernumber=p.customernumber);

23. Join Optimizations:

Map Join (Broadcast Join or Broadcast-Hash Join)



This will be used in the Star schema data model in the DWH applications, small tables (dimension tables) in memory in all of the mappers and big table (fact table) is joined over it in the mapper. This avoids shuffling cost that is inherent in Common-Join. A hash table would be created for each of the small table (dimension table), using join key as the hash table key.

```
set hive.auto.convert.join=true;
set hive.auto.convert.join.noconditionaltask = true;
set hive.auto.convert.join.noconditionaltask.size = 10000000;
//default 10mb
set hive.mapjoin.smalltable.filesize=100;
//default 25mb
```

```
explain select /*+ MAPJOIN(a) */ a.* from customers a inner join payments b on
a.customernumber=b.customernumber;
```

Bucket Map Join

While the tables are large and all the tables used in the join are bucketed on the join columns we use Hive Bucket Map Join feature. Moreover, one table should have buckets in multiples of the number of buckets in another table in this type of join

While all tables are Large.

Also, while all tables are bucketed using the join columns.

While by using the join columns, Sorted.

Also, when the number of buckets is same as the number of all tables.

```
set hive.optimize.bucketmapjoin;
set hive.enforce.bucketing = true ;
```

```
CREATE TABLE `customersbuck` ( `customerNumber` int , `customerName` varchar(50) ,
`contactLastName` varchar(50) , `contactFirstName` varchar(50) , `phone` varchar(50) , `addressLine1`
varchar(50) , `addressLine2` varchar(50) , `city` varchar(50) , `state` varchar(50) , `postalCode`
varchar(15) , `country` varchar(50) , `salesRepEmployeeNumber` int , `creditLimit` decimal(10,2) )
clustered by (customernumber) sorted by (customernumber) into 10 buckets
row format delimited fields terminated by ',';
```

```
insert overwrite table customersbuck select * from customers;
CREATE TABLE `paymentsbuck` ( `customerNumber` int , `checkNumber` varchar(50) , `paymentDate`
date , `amount` decimal(10,2) )
clustered by (customernumber) sorted by (customernumber) into 10 buckets
row format delimited fields terminated by ',';
insert overwrite table paymentsbuck select * from payments;
```

```
explain select /*+ MAPJOIN(a) */ a.* from customersbuck a inner join paymentsbuck b on
a.customernumber=b.customernumber;
```

Sort Merge Bucket Join/

While each mapper reads a bucket from the first table and the corresponding bucket from the second table, in SMB join. Also, when the tables are large we can use Hive Sort Merge Bucket join.

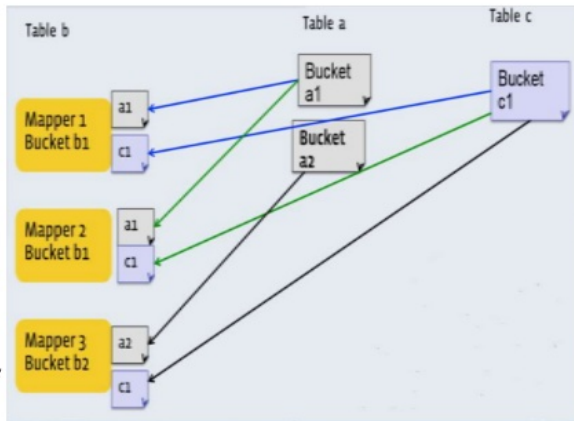
However, using the join columns, all join the columns are bucketed and sorted in SMB. Although, make sure in SMB join all tables should have the same number of buckets.

While all tables are Large.

Also, while all tables are bucketed using the join columns.

While by using the join columns, Sorted.

Also, when the number of buckets is same as the number of all tables.

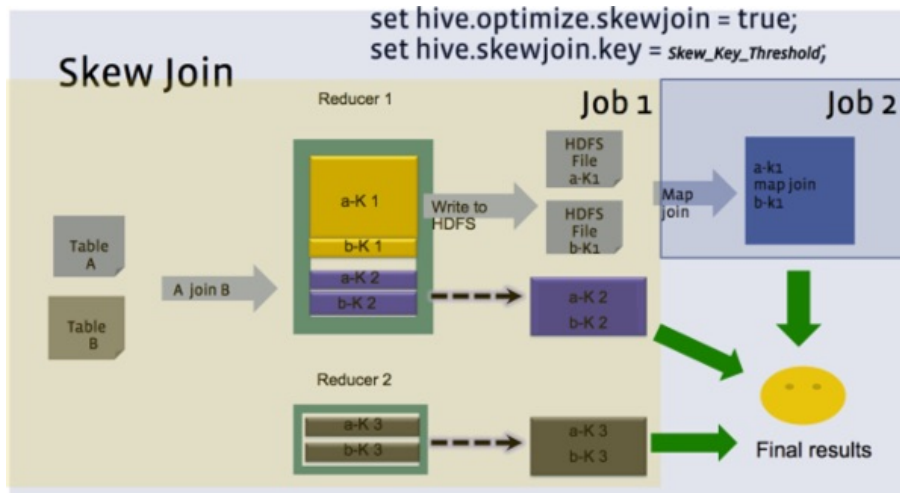


```
set hive.input.format= org.apache.hadoop.hive.ql.io.BucketizedHiveInputFormat;
set hive.auto.convert.sortmerge.join=true;
set hive.optimize.bucketmapjoin = true;
set hive.optimize.bucketmapjoin.sortedmerge = true;
```

```
explain select a.* from customersbuck a inner join paymentsbuck b on
a.customernumber=b.customernumber;
```

3Skew Join :

When there is a table with skew data in the joining column, we use skew join feature. On defining what is skewed table, it is a table that is having values that are present in large numbers in the table compared to other data.



```
set hive.optimize.skewjoin=true;
set hive.skewjoin.key=100000;
```

24. Set operators :

Show the customers made the payment in a range of amounts of 6000 to 10000 and 30000 to 50000.

```
select * from payments where amount between 6000 and 10000
union all select * from payments where amount between 30000 and 50000;
```

```
select customernumber,amount,'NA' from payments where amount between 6000 and 10000
union all select customernumber,0,country from customers where amount between 30000 and 50000;
```

25. show me the combined result of two different tables picking 1 column from table1 and 2 columns from table2.

```
select p.customernumber,p.amount,'NA' as COUNTRY from payments as p where P.amount between 6000 and 10000
union all select customernumber,0 as AMOUNT,country from customers where customernumber>400;
```

26. Customer who made the payment (intersect), who are the anonymous customers minus (above use cases)?

Hive doesn't support intersect or minus, to use it we need to use inner join or left outer join respectively.

27. Show me the customers who have more than 1 phone number and more than 1 address?

```
insert into `customers` values (103,'Atelier graphique','Schmitt','Carine ','908-199-0411','55, rue Royale',NULL,'Nantes',NULL,'44000','France',1370,'21000.00');
```

```
select customernumber,count(distinct phone) as cnt,count(distinct addressline1) as cnt2
from customers
group by customernumber
having cnt>1 and cnt2>1;
```

Complex types:

28. How do you create structure data from set of columns in hive, how do you group the list of columns and form complex type like structure?

```
select named_struct("id",customernumber,"name",customername) from customers;
```

29. How to collect the columns as array or how to group or convert column to rows in Hive? Eg. Group all phone numbers of the given customer? or how do you pivot a column into row?

```
select customernumber,collect_list(phone) as grouped_phone from customers group by
customernumber;
```

Copy the below data into linux some location.

```
112,Customer very concerned about the exact color of the models. There is high risk that he may dispute the order because there is a slight color mismatch,home$about-us$profile$cart$order$exit,1$2$3$4$5$6
114,Can we deliver the new Ford Mustang models by end-of-quarter?,home$about-us$profile$cart$order$exit,1$2$3$4$5$6
124,Custom shipping instructions were sent to warehouse,home$adv-banner117$cart$exit,1$2$3$4
124,They want to reevaluate their terms agreement with Finance.,home$about-us$profile$cart$order$exit,1$2$3$4$5$6
128,Check on availability.,home$about-us$profile$cart$order,1$2$3$4$5
128,Customer very concerned about the exact color of the models. There is high risk that he may dispute the order because there is a slight color mismatch,google-search$billing$exit,1$2$3
131,Difficult to negotiate with customer. We need more marketing materials,home$adv-banner114$cart$exit,1$2$3$4
131,Order was mistakenly placed. The warehouse noticed the lack of documentation.,google-search$billing$exit,1$2$3
131,They want to reevaluate their terms agreement with Finance.,google-search$billing$exit,1$2$3
141, I need all the information I can get on our competitors.,google-search$billing$exit,1$2$3
141,Cautious optimism. We have happy customers here if we can keep them well stocked. I need all the information I can get on the planned shipments of Porches,home$adv-banner114$cart$exit,1$2$3$4
141,Customer doesn't like the colors and precision of the models.,google-search$billing$exit,1$2$3
141,Customer requested that DHL is used for this shipping,home$about-us$profile$cart$exit,1$2$3$4$5
141,Difficult to negotiate with customer. We need more marketing materials,google-search$billing$exit,1$2$3
141,Disputed then Resolved on 3/15/2016. Customer doesn't like the craftsman ship of the models.,home$about-us$profile$cart$order$exit,1$2$3$4$5$6
141,This customer found a better offer from one of our competitors. Will call back to renegotiate.,google-search$billing$exit,1$2$3
144,Customer requested that ad materials (such as posters pamphlets) be included in the shipment,google-search$billing$exit,1$2$3
144,The outstanding balance for this customer exceeds their credit limit. Order will be shipped when a payment is received.,google-search$billing$exit,1$2$3
145,Customer claims container with shipment was damaged during shipping and some items were missing. I am talking to FedEx about this.,home$about-us$profile$cart$order$exit,1$2$3$4$5$6
148,They want to reevaluate their terms agreement with Finance.,google-search$billing$exit,1$2$3
```

148,This order was on hold because customers's credit limit had been exceeded. Order will ship when payment is received,home\$about-us\$profile\$cart\$order\$exit,1\$2\$3\$4\$5\$6

148,We need to keep in close contact with their Marketing VP. He is the decision maker for all their purchases.,home\$about-us\$profile\$cart\$order\$exit,1\$2\$3\$4\$5\$6

151,Customer requested special shippment. The instructions were passed along to the warehouse,home\$exit,1\$2

167,We need to keep in close contact with their Marketing VP. He is the decision maker for all their purchases.,home\$about-us\$profile\$cart\$order\$exit,1\$2\$3\$4\$5\$6

171,Customer has worked with some of our vendors in the past and is aware of their MSRP,home\$about-us\$profile\$cart\$order\$exit,1\$2\$3\$4\$5\$6

172,Customer requested that DHL is used for this shipping,home\$about-us\$profile\$cart\$order\$exit,1\$2\$3\$4\$5\$6

173,Can we deliver the new Ford Mustang models by end-of-quarter?,home\$about-us\$profile\$cart\$order\$exit,1\$2\$3\$4\$5\$6

175,Customer requested that DHL is used for this shipping,home\$about-us\$profile\$cart\$order\$exit,1\$2\$3\$4\$5\$6

186,Cautious optimism. We have happy customers here if we can keep them well stocked. I need all the information I can get on the planned shipments of Porches,home\$adv-banner117\$cart\$exit,1\$2\$3\$4

189,We must be cautions with this customer. Their VP of Sales resigned. Company may be heading down.,home\$adv-banner117\$cart\$exit,1\$2\$3\$4

201,Customer disputed the order and we agreed to cancel it. We must be more cautions with this customer going forward since they are very hard to please. We must cover the shipping fees.,home\$adv-banner117\$cart\$exit,1\$2\$3\$4

202,Can we renegotiate this one?,home\$adv-banner117\$cart\$exit,1\$2\$3\$4

202,Customer requested that FedEx Ground is used for this shipping,home\$adv-banner117\$cart\$exit,1\$2\$3\$4

205,They want to reevaluate their terms agreement with Finance.,home\$adv-banner117\$cart\$exit,1\$2\$3\$4

205,This order was disputed and resolved on 2/1/2016. Customer claimed that container with shipment was damaged. FedEx's investigation proved this wrong.,home\$adv-banner117\$cart\$exit,1\$2\$3\$4

216,Customer has worked with some of our vendors in the past and is aware of their MSRP,home\$adv-banner117\$cart\$exit,1\$2\$3\$4

216,Customer inquired about remote controlled models and gold models.,home\$adv-banner117\$cart\$exit,1\$2\$3\$4

216,Customer is interested in buying more Ferrari models,home\$adv-banner117\$cart\$exit,1\$2\$3\$4

240,Customer requested that ad materials (such as posters pamphlets) be included in the shippment,home\$adv-banner117\$cart\$exit,1\$2\$3\$4

242,Custom shipping instructions sent to warehouse,home\$adv-banner117\$cart\$exit,1\$2\$3\$4

242,Customer is interested in buying more Ferrari models,home\$adv-banner117\$cart\$exit,1\$2\$3\$4

250,We must be cautions with this customer. Their VP of Sales resigned. Company may be heading down.,home\$adv-banner117\$cart\$exit,1\$2\$3\$4

259,We must be cautions with this customer. Their VP of Sales resigned. Company may be heading down.,home\$adv-banner117\$cart\$exit,1\$2\$3\$4

276,They want to reevaluate their terms agreement with Finance.,home\$adv-banner117\$cart\$exit,1\$2\$3\$4

278,Cautious optimism. We have happy customers here if we can keep them well stocked. I need all the information I can get on the planned shipments of Porches,home\$adv-banner117\$cart\$exit,1\$2\$3\$4

278,Customer very concerned about the exact color of the models. There is high risk that he may dispute the order because there is a slight color mismatch,home\$adv-banner117\$cart\$exit,1\$2\$3\$4

282,Can we renegotiate this one?,home\$adv-banner117\$cart\$exit,1\$2\$3\$4

286,Check on availability.,home\$adv-banner117\$cart\$exit,1\$2\$3\$4

```
create table orderpages (customernumber string,comments string,pagenavigation array
<string>,pagenavigationidx array <int>)
row format delimited fields terminated by ','
collection items terminated by '$';
```

load data local inpath '/home/hduser/orderstg' overwrite into table orderpages;

30. How to collect the array as rows or Eg. UnGroup all webpage visited numbers of the given customer? or how do you unpivot an array column into row of values?
select explode(pagenavigation) from orderpages ;

31. Show me the position of the array elements in an array column in Hive? or show the order in which customer navigated in the webpage which is stored as array?
select posexplode(pagenavigation) from orderpages ;

32. Display the customer number, comments and the pages navigated from the array data ?

We need to use Lateral view is used in conjunction with user-defined table generating functions such as explode() UDTF (user defined table generating functions). A lateral view first applies the UDTF to each row of base table and then joins resulting output rows to the input rows to form a virtual table having the supplied table alias.

```
select customernumber,comments,pgnavigation_column
from orderpages lateral view explode(pagenavigation) exploded_tbl as pgnavigation_column;
```

Ordering:

33. Why order by in hive is more costly? how can you avoid using it?

Order by will go with single reducer operation, can be avoided with multi reducer operation using distribute by and sort by.

```
set mapred.reduce.tasks=3;
select * from payments order by customernumber,amount ;
select * from payments distribute by customernumber sort by customernumber,amount ;
select * from payments cluster by customernumber,amount;
```

or

```
select * from payments order by customernumber,amount limit 100;
```

tn240511ka441122