geolocation_zip_code  xamining the sec s.head()  06b8999e2fba1a1fbc8 18955e83d337fd6b2d 4e7b3e00288586ebd08 b2b6027bc5c5109e529 4f2d8ab171c80ec836  xamining the thi l.head()	1037 -23. 1046 -23.	\\sujoydutta sujoydutta\\? rs\\sujoydut	\\Desktop\ Desktop\\D ta\\Deskto			atasets fotasets for Sets for Datasets	r ML\\EDA\\ ML\\EDA\\OF for ML\\EDA	EO_LOCATION.c  JSTOMERS.csv'  ELLERS.csv')  \ORDERS.csv'  \ORDER_ITEMS.  RDER_PAYMENTS  \\ORDER_REVI  PRODUCTS.csv'	) csv') .csv') EW_RATIN
06b8999e2fba1a1fbc8 18955e83d337fd6b2d 4e7b3e00288586ebd08 b2b6027bc5c5109e529 4f2d8ab171c80ec836	1035 -23. 1012 -23.	545621 -4 546081 -4 544392 -4 541578 -4	ation_lng ge .6.639292 .6.644820 .6.639499 .6.641607 .6.635361	Akkarampal Akkarampal Akkarampal Akkarampal Akkarampal	le Andhra le Andhra le Andhra le Andhra	Pradesh Pradesh Pradesh Pradesh Pradesh			
I.lieau()	ef6b18a428ac77 8712fdd0374a03 9d4dc6358b12c3 4f7c12e35b23ad	861eff4711a542 290c77bc529b7a 060e732b5b29e8 259dac757896c 345ecd01c38d18	2e4b93843c6dd ac935b93aa66d 181a18229c7b I24d7702b9acI	d7febb0 c3333dc3 o0b2b5e bbff3f3c	stomer_zip_co	14409 9790 1151 8775 13056	Adilabad Adoni Akkarampalle	customer_state Andhra Pradesh Andhra Pradesh Andhra Pradesh Andhra Pradesh Andhra Pradesh	
3442f8959a84dea7ee d1b65fc7debc3361ea6 ce3ad9de960102d0677 c0f3eea2e14555b6fae 51a04a8a6bdcb23dec xamining the founders.head()	197c632cb2df15 86b5f14c68d2e2 7a81f5d0bb7b2d eea3dd58c1b1c3	·	13023 13023 20031 Ba 4195 Akkara		dhra Pradesh				
e481f51cbdc54678b7 53cdb2fc8bc7dce0b67- 47770eb9100c2d0c449 949d5b44dbf5de918f- ad21c59c0840e6cb83a xamining the fifems.head()	41e2150273451 946d9cf07ec65d e9c16f97b45f8a a9ceb5573f8159	9ef432eb6251297 b0830fb4747a6c6 41ce2a54c0b03b f88197465ea7920 8ab97904e6daea8	304e76186b10 5d20dea0b8c8 f3443c3d931a3 adcdbec73753	0a928d 02d7ef 367089 364d82	delivered delivered delivered delivered delivered	10, 7/2	e_timestamp /2/2017 10:56 24/2018 20:41 8/8/2018 8:38 18/2017 19:28 13/2018 21:18	order_approved_a 10/2/2017 11:0 7/26/2018 3:2 8/8/2018 8:5 11/18/2017 19:4 2/13/2018 22:2	07 24 55
00010242fe8c5a6d1ba 00018f77f2f0320c557 000229ec398224ef6ca 00024acbcdf0a6daa1e 00042b26cf59d7ce69d xamining the six y.head()	2dd792cb16214 190d7a144bdd3 a0657da4fc703e a931b038114c75 dfabb4e55b4fd9	1 1 1	4244733e06e70 e5f2d52b8021 c777355d18b <sup>-1</sup> 7634da152a4 ac6c3623068f3	ecb4970a6e2 89ee658865c 72b67abbeef 4610f1595efa	a93d83a8f c 9df44fd0fd 5 32f14722fc 9	ld7ddc04e1l 5b51032eddd d7a1d34a50	Bac8b2bce089ec b6c2c614352b38 d242adc84c38ac 5240900642527 51e74553ab940	2a041202 9 33efe2d36 9 cab88f23d 1/ 5ba1c2b4 8/	ing_limit_d 9/19/2017 9 5/3/2017 11 (18/2018 14 (15/2018 10 (13/2017 13
b81ef226f3fe1789b16 a9810da82917af2d9 25e8ea4e93396b6fa0c ba78997921bbcdc1373 42fdf880ba16b47b592 xamining the sev tings.head()	e8b2acac839d17 aefd1278f1dcfa0 d3dd708e76c1bd Bbb41e913ab953 251dd489d4441a	payment_seque	1 cred 1 cred 1 cred 1 cred 1 cred	dit_card dit_card dit_card dit_card dit_card		8 1 1 8 2	99.33 24.39 65.71 107.78 128.45	ny answer timest	amn
7bc2406110b926393aa 80e641a11e56f04c1aa 228ce5500dc1d8e020c e64fb393e7b32834bb f7c4243c7fe1938f181 xamining the eigod.head()	d469d5645fdfde d8d1322874b6f0 789ff8bb30750e bec41a392bdeb	8e6bfb81e283f	9712e6da79b0 7796b98fdf73d f2ecdecbb34b 9be170737859 fa7e4f11123a3	beba33 ed034b d3511b fb894f1	4 5 5 5 5	1/18/2 3/10/2 2/17/2 4/21/2 3/1/2	018 0:00 018 0:00 018 0:00 017 0:00 018 0:00	2/18/2018 1 4/21/2018 1 3/11/2018 2 2/18/2018 1 4/21/2017 2 3/2/2018 1	1:46 3:05 4:36 2:02 0:26
1e9e8ef04dbcff4541e 3aa071139cb16b67cat 96bd76ec8810374ed1b cef67bcfe19066a932b 9dc1a7de274444849c  erging the datas s_ord= pd.merge( s_ord.head()	ed26657ea517e5 9e5dea641aaa2f b65e291975717f b7673e239eb23d c219cff195d0b71	Pe Sports Hou	Art S_Leisure Baby ssewares		40.0 44.0 46.0 27.0 37.0		287.0 276.0 250.0 261.0 402.0	1.0 1.0 1.0 1.0 4.0	
18955e83d337fd6b2d 4e7b3e00288586ebd08 b2b6027bc5c5109e529 4f2d8ab171c80ec8364 erging the datas	ef6b18a428ac77 8712fdd0374a03 9d4dc6358b12c3 4f7c12e35b23ad	290c77bc529b7a 060e732b5b29e8 259dac757896c 345ecd01c38d18 and items	2e4b93843c6dd ac935b93aa66d 181a18229c7b 124d7702b9acl a9036ed96c73	d7febb0 c333dc3 o0b2b5e bbff3f3c 8b8d066		14409 9790 1151 8775 13056	Adilabad Adoni Akkarampalle	Andhra Pradesh Andhra Pradesh Andhra Pradesh	00e7ee1b 2915012 b2059ede 9516709 6b7d50
18955e83d337fd6b2d 4e7b3e00288586ebd08 b2b6027bc5c5109e529 4f2d8ab171c80ec836 erging the datas s_ord_it_sel= pd	ef6b18a428ac77 8712fdd0374a03 9d4dc6358b12c3 4f7c12e35b23ad sets cus_ord_1	290c77bc529b7a 060e732b5b29e8 259dac757896c 345ecd01c38d18 it and selle	e4b93843c6dd ac935b93aa66d 181a18229c7b 124d7702b9acl a9036ed96c73	d7febb0 c3333dc3 o0b2b5e bbff3f3c 8b8d066		14409 9790 1151 8775 13056	Adilabad Adoni Akkarampalle	Andhra Pradesh Andhra Pradesh Andhra Pradesh	00e7ee1b 2915012 b2059ed 951670 6b7d50
4e7b3e00288586ebd08 b2b6027bc5c5109e529 4c06b42fbf7b97ab107	8712fdd0374a03 9d4dc6358b12c3 779cda5549cd1c	060e732b5b29e8 259dac757896c 07d190f123147d	2e4b93843c6dd 181a18229c7b 124d7702b9acl	d7febb0 n0b2b5e bbff3f3c	stomer_zip_co	14409	Adilabad  Akkarampalle  Akkayapalle	Andhra Pradesh  Andhra Pradesh  Andhra Pradesh	
s_ord_it_sel_it= s_ord_it_sel_it. 06b8999e2fba1a1fbc8	customer_id	s_ord_it_sel 861eff4711a542	customer_un	<b>iique_id cu</b> d7febb0			<b>customer_city</b> Adilabad	customer_state  Andhra Pradesh	00e7ee1lb2059ed
4c06b42fbf7b97ab107 8276de07ef25225d412 bws × 26 columns erging the datas s_ord_it_sel_it_	779cda5549cd1c 2b8462d73f8664 sets cus_ord_1 pay= pd.merge	07d190f123147d 332cf4e83e1600 it_sel_it_an	9e89d4b92254 94ba7dca932ce d pay	13d7948 e82475b	order_id']	8775 65075 90010	Chilakaluripet Eluru		951670 997fd4b b63fff1
4e7b3e00288586ebd08 b2b6027bc5c5109e529 4c06b42fbf7b97ab107	8712fdd0374a03 9d4dc6358b12c3 779cda5549cd1c	060e732b5b29e8 259dac757896c 07d190f123147d	2e4b93843c6dd 181a18229c7b 124d7702b9acl	d7febb0 n0b2b5e bbff3f3c	stomer_zip_co	14409 1151 8775 65075 90010	Adilabad  Akkarampalle  Akkayapalle	Andhra Pradesh  Andhra Pradesh  Andhra Pradesh	
s_ord_it_sel_it_ s_ord_it_sel_it_ 06b8999e2fba1a1fbc8 4e7b3e00288586ebd08 b2b6027bc5c5109e529	customer_id 88172c00ba8bc7 8712fdd0374a03	861eff4711a542 060e732b5b29e8	d_it_sel_i customer_un 2e4b93843c6dd 181a18229c7b	t_pay, ra  ique_id custod7febb0  i00b2b5e  bbff3f3c		14409 1151 8775	customer_city  Adilabad  Akkarampalle  Akkayapalle	customer_state Andhra Pradesh Andhra Pradesh Andhra Pradesh	00e7ee1lb2059ed
8276de07ef25225d412  bws × 34 columns  erging the datas  nal= pd.merge(cu  nal.head()  06b8999e2fba1a1fbc88	customer_id	332cf4e83e1600 it_sel_it_pa _it_pay_rat, 861eff4711a542e	y and prod prod, lef :ustomer_unic	e82475b  luct t_on='pro que_id cust 7febb0		90010 , right_	Eluru on='product	Andhra Pradesh  a_id', how='i  customer_state  Andhra Pradesh  Jammu &	
8912fc0c3bbf1e2fbf35 8912fc0c3bbf1e2fbf35 ows × 43 columns	5819e21706718 9 5819e21706718 9	eae34bbd3a474e	c5d07949ca7d	le67c0		68030 68030 68030	Rehambal Rehambal	Jammu & Kashmir Jammu & Kashmir	c1d2b34fe c1d2b34fe c1d2b34fe
t64Index: 160171 ta columns (tota	entries, 0 t 1 43 columns)  ue_id code_prefix  e  e_timestamp d_at ed_carrier_da	Non-Nu:  Non-Nu:   160171  160171  160171  160171  160171  160171  160171  160171  160154	non-null non-null non-null non-null non-null non-null non-null non-null non-null	object int64 object					
2 order_item_id 3 product_id_x 4 seller_id 5 shipping_limi 6 price_x 7 freight_value 8 seller_zip_co 9 seller_city 0 seller_state 1 order_item_id 2 product_id_y 3 shipping_limi 4 price_y 5 freight_value	t_date_x  x de_prefix y t_date_y	160171 160171 160171 160171 160171 160171 157950 157950 160171 160171 160171	non-null	int64 object object float64 float64 int64 object object int64 object float64 float64					
payment_value review_id review_score review_creati review_answer product_id product_categ product_name_ product_descr product_photo product_lengt product_heigh	on_date _timestamp ory_name lenght iption_lenght s_qty t_g h_cm t_cm	160171 160171 160171 160171 160171 160171 157924 157953 157953 157953 160147 160147	non-null	int64 float64 object object object object float64 float64 float64 float64 float64 float64					
Removing redunda lumns_to_drop = nal.drop(columns  eeing the final nal.dtypes stomer_id stomer_unique_id stomer_zip_code_s stomer_city stomer_state	+ MB  ant columns [col for col =columns_to_c	in final.co drop, inplace mns  object object int64 object object		ol.endswi	th("_y")]				
der_approved_at der_delivered_ca der_delivered_cu der_estimated_de der_item_id_x oduct_id_x ller_id ipping_limit_dat ice_x eight_value_x ller_zip_code_pr ller_city ller_state	rrier_date stomer_date livery_date  e_x efix	object object object object object int64 object object object float64 int64 object							
yment_type yment_installmen yment_value view_id view_score view_creation_da view_answer_time oduct_id oduct_category_n oduct_name_lengh oduct_descriptio	te stamp ame t n_lenght	object int64 float64 object int64 object object object object float64 float64 float64 float64 float64							
ype: object  enaming the columnal.rename(columnal.head()  06b8999e2fba1a1fbc88	customer_id 8172c00ba8bc7	cem_id_x': ' 861eff4711a542e eae34bbd3a474e	<b>customer_unic</b> 4b93843c6dd7 c5d07949ca7d	que_id cust 7febb0 le67c0			ustomer_city	Andhra Pradesh  Jammu &  Kashmir	
8912fc0c3bbf1e2fbf35  ows × 38 columns  onverting to cor nal['order_purch nal['order_appro nal['order_deliv nal['order_deliv	5819e21706718 9  Frect format  Lase_timestamp  Eved_at'] = po  Frered_carrier_  Frered_customer	p'] = pd.to_d.to_datetim_date'] = pd	c5d07949ca7d  datetime(fe(final['o .to_datetid.to_datet	inal['ord rder_appr me(final[ ime(final	roved_at']) 'order_de] ['order_de	_ livered_c elivered_	arrier_date customer_da	Kashmir  Jammu & Kashmir	c1d2b34fe
nal['review_crea nal['review_answ nal['shipping_li "xamining the col nal.info()  lass 'pandas.cor t64Index: 160171 ta columns (tota	tion_date'] =  ver_timestamp' mit_date'] =  tumns  e.frame.DataF entries, 0 t 1 38 columns)	rame'> Non-Nu: 160171 160171	time(final atetime(fi ime(final [ Ime (final [ Ime (Ime (Ime (Ime (Ime (Ime (Ime (Ime	['review_nal['review_a	creation_c ew_answer_	date']) _timestam			
customer_stat order_id order_status order_purchas order_approve order_deliver 0 order_deliver 1 order_estimat 2 order_item_id 3 product_id 4 seller_id 5 shipping_limi 6 price 7 freight_value	e_timestamp d_at ed_carrier_da ed_customer_d ed_delivery_d	160171 160171 160171 160171 160154 158239 date 156691 date 160171 160171 160171 160171 160171	non-null	object object datetime datetime datetime datetime datetime datetime int64 object object datetime float64 float64	64[ns] 64[ns] 64[ns] 64[ns]				
seller_state payment_seque payment_type payment_insta payment_value review_id review_score review_creati review_answer product_id product_categ product_name_ product_descr product_photo	ntial  llments  on_date _timestamp  ory_name lenght iption_lenght s_qty	157950 160171 160171 160171 160171 160171 160171 160171 160171 157924 157953 157953	non-null	object int64 object int64 float64 datetime object object float64 float64 float64					
6 product_heigh 7 product_width ypes: datetime64 mory usage: 47.7 ropping null val nal=final.dropna nal.info() lass 'pandas.cor t64Index: 152451	t_cm _cm [ns](8), floa + MB wes () e.frame.DataF entries, 0 t	160147 160147 at64(10), ins	non-null non-null t64(6), ob	float64 float64 ject(14)					
customer_zip_ customer_city customer_stat order_id order_status order_purchas order_approve order_deliver order_deliver order_estimat order_item_id product_id seller_id	code_prefix  e  e_timestamp d_at ed_carrier_da ed_customer_d ed_delivery_d	152451 152451 152451 152451 152451 152451 152451 14te 152451 14te 152451 152451 152451 152451 152451	non-null	int64 object object object datetime	64[ns] 64[ns] 64[ns] 64[ns]				
7 freight_value 8 seller_zip_co 9 seller_city 0 seller_state 1 payment_seque 2 payment_insta 4 payment_value 5 review_id 6 review_score 7 review_creati 8 review_answer 9 product_id 0 product_categ	de_prefix ntial llments on_date _timestamp ory_name	152451 152451 152451 152451 152451 152451 152451 152451 152451 152451 152451 152451	non-null	float64 int64 object int64 object int64 float64 object int64 datetime datetime object object					
3 product_photo 4 product_weigh 5 product_lengt 6 product_heigh 7 product_width ypes: datetime64 mory usage: 45.4  Removing duplica nal = final.loc[ nal.head()	s_qty t_g h_cm t_cm cm [ns](8), floa + MB  te columns :, ~final.col	152451 152451 152451 152451 152451 at64(10), in	non-null non-null non-null non-null t64(6), ob	float64 float64 float64 float64 ject(14)					00.27.0.11.0
8912fc0c3bbf1e2fbf35 8912fc0c3bbf1e2fbf35 8912fc0c3bbf1e2fbf35 bws × 37 columns	5819e21706718 9 5819e21706718 9 5819e21706718 9	eae34bbd3a474e eae34bbd3a474e	c5d07949ca7d c5d07949ca7d	le67c0 le67c0		68030 68030 68030 68030	Rehambal Rehambal Rehambal	Jammu & Kashmir  Jammu & Kashmir  Jammu & Kashmir	c1d2b34fe c1d2b34fe c1d2b34fe c1d2b34fe
tal_revenue = fi int(f'Total Reve tal Revenue: INR Calculating Tota tal_quantity = f int(f'Total Quan tal Quantity of f Calculating Tota tal_products = f int(f'Total Prod	nal['price']. nue: INR {tot  16595198.70  al Quantity of  inal['order_i  tity of products sold  al Products of  inal['products  lucts listed:	f products sitem_id'].suracts sold: {  1: 247992  f products 1  2_id'].nuniq	old m() total_quan isted ue()	tity}')					
tal_categories = int(f'Total Cate  tal Categories a  Calculating Tota  tal_locations = int(f'Total Loca  tal Locations de  Calculating New w_customers_per_ int('New Custome	efinal['productions available: 71  Al Locations of final['custom ations deliver livered to: 4  Customers Acquired Ers Acquired E	delivered to mer_city'].n red to: {total	_name'].nu _categorie unique() al_locatio  Month nal['order	nique() s}') ns}')	_timestamr	o'].dt.tc	_period("M'	'))['customer	_unique
w Customers Acqu der_purchase_tim 16-10	ired Every Mc	onth:							
Popular Products	by Month			er purcha	se timesta	amo'l.dt.	to period(	'M"))['mroduc	t id'l.
<pre>pular_products_b int('Popular Pro int(popular_prod pular Products b</pre>	oy_month = pop oducts by Month ducts_by_month y Month: timestamp pr 2016-10 2016-12 2017-01 2017-02 2017-03 2017-04 2017-05 2017-06	ch:')  coduct_id  483  1 1669 2399 3998 3222 5341 4504							
Popular Categori	2017-07 2017-08 2017-09 2017-10 2017-11 2017-12 2018-01 2018-02 2018-03 2018-04 2018-05 2018-06 2018-07 2018-08	6334 6879 6620 7600 12177 8193 10770 10982 10974 10967 11009 9847 9296 9186							
pular_categories pular_categories pular_categories pular_categories  customer_state  13 Andhra Pradesh  14 Andhra Pradesh  15 Andhra Pradesh  24 Andhra Pradesh	s_by_state_mor s_by_state_mor s_by_state_mor	mestamp prod 2016-10 2016-10 2016-10 2016-10 2016-10	r_categori  uct_category_  Furniture_  Market  Perf	name orde Decor Place Tumery Beauty Toys	er_item_id  89  42  38  30  19				
West Bengal West Bengal West Bengal West Bengal West Bengal West Bengal Frows × 4 columns West Bengal Frows × 4 columns	10 Most Exper	2018-08 2018-08 2018-08 2018-08 2018-08	Home_C Pet Stat Watches	iances confort c_Shop ionery s_Gifts	1 1 1 1	<b>lse</b> ).head	.(10)		
6319 69c590f7ff 1232 1bdf5e6731 3532 a6492cc693 3525 c3ed642d59 5955 259037a6a4 509 a1beef8f39 5258 6cdf8fc1d7 3956 dd113cb02b 444 6902c1962d	ive Products:  products:  8f021502940f2 c7bf8db97190b 585cf01aa8169 76c469ab6f61d 2594bb648ff4a 1845e455183f8 92dbd4cd87267 41c76586d8b6b 2af9c8e5787e8 d19d540807d0a	product_id 251d4cce7f 66cb6ed62e 0c7028d6ad 88f44de961 404cee2747 89c5035f18 796aa69c53 615e9eef30	price 6735.00 6729.00 6499.00 4799.00 4690.00 4590.00 4399.87 4099.99	e']])					
'Unique Cu 'Customer	<pre>comer_retention = pd.to_datet er_month = dat ution = n_cust</pre>	cime(data[monta.groupby('n	nth_column month')['c onth.pct_c ex, r_month.va	ustomer_ustomer_ushange() *	_ period("M' nique_id']	')			
return retenti  evenue by custom f calculate_reve  data['customer data['customer data['customer data['customer data['customer revenues_by_cu return revenue	enues_by_custo e_type'] = dat e_type'] = dat e_type'] = dat e_type'] = dat evenues from estomer_type = estomer_type =	ca.groupby('ca['order_puca['customer]' 'New' and 'E cata.group	customer_u rchase_tim _type'].re xisting' c by(['order	estamp'] place({Tr rustomers _purchase	== data['crue: 'New', each month timestamp	customer_ False:	<pre>type'] 'Existing') omer_type']</pre>	))['price'].s	um().un
ustomer Retentio	on on a Month- a = calculate_ etention on a ention)	-to-Month Ba customer_re Month-to-Mo	tention(finth Basis:	') ) N 2 0 6 6 8					
2016-10 2016-12 2017-01 2017-02 2017-03 2017-04	2961 3696 3987		-11.37384 24.82269 7.87337 -0.70228 8.56276 62.26151 -24.41927 27.41415 -7.10244 7.29283 -1.62832 0.22779 -9.51515 -0.06697 3.53552	0 5 7 2 8 7 2 3 2 5 4 0 2 9					
int('Customer Re int(customer_ret  stomer Retention	3959 4298 6974 5271 6716 6239 6694 6585 6600 5972 5968 6179		nues_by_cu		is				
int('Customer Re int(customer_ret  stomer Retention	4298 6974 5271 6716 6239 6694 6585 6600 5972 5968 6179 **Sting/New Customer_type = callon	Coulate_reverse		onth Basi	<pre>rpe(final) th Basis:' s:</pre>	1 1 1 1 1 1			
int('Customer Re int(customer_ret  stomer Retention	4298 6974 5271 6716 6239 6694 6585 6600 5972 5968 6179  Sting/New Cust mer_type = call com Existing/N customer_type ting/New Cust r_purchase_ti 2016-10-03 1 2016-10-03 2 2016-10-03 2 2016-10-03 2 2018-08-29 1 2018-08-29 1 2018-08-29 1 2018-08-29 1	Culate_reverse		onth Basi	rpe(final) th Basis:' s: Customers Nan Nan Nan Nan Nan				
int('Customer Re int(customer_ret  stomer Retention	4298 6974 5271 6716 6239 6694 6585 6600 5972 5968 6179  **Sting/New Cust er_type = call com Existing/N customer_type ting/New Cust r_purchase_ti 2016-10-03 0 2016-10-03 1 2016-10-03 2 2016-10-03 2 2018-08-29 1 2018-08-29 1 2018-08-29 1 2018-08-29 1 2018-08-29 1 2018-08-29 1 2018-08-29 1 2018-08-29 1 2018-08-29 1 2018-08-29 1 2018-08-29 1	Culate_reve.  New Customer  Comers on a lance of the comers on a lance of the comers on a lance of the comers of t	omer_id').  gories  x, es	onth Basi Existing	rpe(final) th Basis:' s: Customers Nan Nan Nan Nan Nan Nan Nan Nan		'order_item	n_id'] * grou	p['pric
<pre>int('Customer Re int (customer_ret  stomer Retention</pre>	4298 6974 5271 6716 6239 6694 6585 6600 5972 5968 6179  **Sting/New Cust er_type = call com Existing/N customer_type ting/New Cust r_purchase_ti 2016-10-03 0 2016-10-03 1 2016-10-03 2 2016-10-03 2 2016-10-03 2 2018-08-29 1	Culate_reverse   Comers on a land   Comers on a l	omer_id').  gories  x, es  l ow Spendin oney_spent nt']]) customer_s Low Sp High Sp High Sp	onth Basi Existing  g', 'Mode '], q=qua  egment ending ending ending ending ending ending ending	rpe(final) th Basis:  S: Customers Nam	ding', 'M	id-High Spe	ending', 'Hig	
int('Customer_Reint (customer_ret  int (customer_ret  stomer Retention	### 4298 ### 6974 ### 5271 ### 6716 ### 6239 ### 6585 ### 6600 ### 5972 ### 5968 ### 6179  ### 5271 ### 6716 ### 6239 ### 6585 ### 6600 ### 5972 ### 5968 ### 6179  ### 5271 ### 6694 ### 6585 ### 6600 ### 5972 ### 5968 ### 6179  ### 52968 ### 6179  #### 6179  #### 6179  #### 6179  #### 6179  #### 6179  #### 6179  #### 6179  #### 6179  #### 6179  #### 6179  #### 6179  #### 6179  #### 6179  #### 6179  #### 6179  #### 6179  ##### 6179  ##### 6179  ##### 6179  ###########  ###############  #######	Culate_reverse   Commers on a land   Commers	omer_id').  gories  x, es  l ow Spendin oney_spent nt']]) customer_s Low Sp High Sp High Sp High Sp id-High Sp Low Sp High Sp Low Sp High Sp Low Sp Low Sp Low Sp Low Sp Low Sp Low Sp	onth Basi Existing  g', 'Mode '], q=qua  egment ending	pe(final) th Basis:  S: Customers Nan	ding', 'Mabels=cus	id-High Spe	ending', 'Hig	h Spend
int ('Customer Re int (customer_ret int (customer_ret stomer Retention	### ### ### ### ### ### ### ### ### ##	Culate_reverse   Commers on a land   Commers	omer_id').  gories  x, es  l ow Spendin oney_spent nt']]) customer_s Low Sp High Sp High Sp High Sp High Sp Low Sp	g', 'Modera apply(lam egment ending	pe(final) th Basis:  S: Customers Nan	ding', 'Mid-	did-High Spectomer_segments  er_item_id	ending', 'Highents)	h Spend
int ('Customer Re int (customer_ret int (customer_ret stomer Retention	4298 6974 5271 6716 6239 6694 6585 6600 5972 5968 6179  Sting/New Cust resting/New Cust res	Coulate reverse   Coulate re	omer_id').  gories  x, es  l ow Spendin oney_spent nt']]) customer_s Low Sp High Sp High Sp High Sp Low Sp High Sp Low Sp High Sp I how Sp High Sp Low Sp	onth Basi Existing in Existing apply(lam  g', 'Mode '], q=qua  egment ending en	pe(final) th Basis:  S: Customers Nan	ding', 'Mid-	did-High Spectomer_segments  er_item_id	ending', 'Highents)	h Spend
int ('Customer Re int (customer_ret int (customer_se int (data[['customer_se int (data[['selle	### ### ### ### ### ### ### ### ### ##	Coulate reverse   New Customer     New	omer_id').  gories  x, es  l ow Spendin oney_spent nt']]) customer_s Low Sp High Sp High Sp High Sp Low Sp Low Sp  id').appl  ies  s  l w earning' ey_earned' l) seller_seg High ear d-High ear d-High ear d-High ear d-High ear lighest ear High ear heir count r_id'): tolist()  l): [i]), str(  = 1  1 count.item	onth Basi Existing in Existing	pe(final) th Basis:  Customers Nan	ding', 'Mid- coup['ord coup['ord coup['ord	er_item_id'  High earnir omer_segmer	ending', 'Highents)  '] * group['p  ng', 'High ea  nts)	h Spend
	1895 Sedential State of the control	March   Marc	March   Marc	Martin   M		Manufactor   Man	Mate	Martin   M	Martin   M

In [97]:	Payment channels used by most customers:  credit_card 110972 UPI 32160  voucher 7252 debit_card 2067 Name: payment_type, dtype: int64  #Top 10 maximum rated categories top_10_max_rated_categories = final.groupby('product_category_name')['review_score'].mean().nlargest(10) print('Top 10 maximum rated categories:') print(top_10_max_rated_categories)  Top 10 maximum rated categories: product_category_name
T	product_category_name Fashion_Childrens_Clothes 5.000000 Books_Imported 4.626667 Cds_Dvds_Musicals 4.611111 Small_Appliances_Home_Oven_And_Coffee 4.467532 Books_General_Interest 4.461069 Cine_Photo 4.355556 Costruction_Tools_Tools 4.316239 Luggage_Accessories 4.309819 Books_Technical 4.307420 Fashion_Shoes 4.271386 Name: review_score, dtype: float64
[98]:	<pre>#Top 10 minimum rated categories top_10_min_rated_categories = final.groupby('product_category_name')['review_score'].mean().nsmallest(10) print('Top 10 minimum rated categories:') print(top_10_min_rated_categories)  Top 10 minimum rated categories: product_category_name Security_And_Services</pre>
In [99]:	Fashion_Male_Clothing 3.505319 Party_Supplies 3.547170 Christmas_Supplies 3.673387 Name: review_score, dtype: float64  # top 10 products maximum rated  top_10_max_rated_products = final.groupby('product_id', as_index=False)['review_score'].mean().nlargest(10, 're  print('Top 10 maximum rated products:') print(top_10_max_rated_products)
	Top 10 maximum rated products:
In [100	# top 10 products minimum rated top_10_min_rated_products = final.groupby('product_id')['review_score'].mean().nsmallest(10) print('Top 10 minimum rated products:') print(top_10_min_rated_products)  Top 10 minimum rated products: product_id 0009406fd7479715e4bef61dd91f2462
In [101 Out[101]:	<pre>average_rating_by_location = final.groupby('customer_city')['review_score'].mean() average_rating_by_location  customer_city Aadityana</pre>
In [102	Aantaliya 3.724138 Aarambhada 3.205882 Abiramam 5.000000  Zamin Uthukuli 4.000000 Ziauddin Pur 4.172897 Zira 4.600000 Zirakpur 4.500000 Ziro 3.808824 Name: review_score, Length: 4050, dtype: float64  #Average rating by seller average_rating_by_seller = final.groupby('seller_id')['review_score'].mean()
Out[102]:	<pre>average_rating_by_seller  seller_id 0015a82c2db000af6aaaf3ae2ecb0532</pre>
In [103 Out[103]:	<pre>Name: review_score, Length: 2857, dtype: float64  #Average rating by product average_rating_by_product = final.groupby('product_id')['review_score'].mean() average_rating_by_product</pre>
In [104 Out[104]:	<pre>fffe9eeff12fcbd74a2f2b007dde0c58     4.0 Name: review_score, Length: 31083, dtype: float64  #Average rating by category average_rating_by_category = final.groupby('product_category_name')['review_score'].mean() average_rating_by_category</pre>
In [105 Out[105]:	Stationery 4.111327 Tablets_Printing_Image 4.010417 Telephony 3.789723 Toys 4.144156 Watches_Gifts 3.918832 Name: review_score, Length: 71, dtype: float64  #Average rating by month average_rating_by_month = final.groupby(final['order_purchase_timestamp'].dt.to_period("M"))['review_score'].me average_rating_by_month order_purchase_timestamp 2016-10 3.780538
	2016-12 5.000000 2017-01 4.156980 2017-02 4.160900 2017-03 4.023762 2017-04 4.024209 2017-05 4.132185 2017-06 4.133659 2017-07 3.873698 2017-08 4.039831 2017-09 3.993505 2017-10 3.879211 2017-11 3.752484 2017-12 3.799585
	2017-12 3.799985 2018-01 3.766945 2018-02 3.440903 2018-03 3.626481 2018-04 3.950488 2018-05 4.036425 2018-06 3.995735 2018-07 4.187500 2018-08 4.153168 Freq: M, Name: review_score, dtype: float64