Когортный анализ и прогноз товарооборота на январь 2018 года

1. Определим когорту для каждого пользователя, который когда-либо совершал заказ.

```
CREATE TABLE user_cog
SELECT user_id, DATE_FORMAT(MIN(o_date), '%y%m') as cog
FROM main
GROUP BY user_id;
```

2. Получим суммарные покупки каждой когорты по каждому месяцу.

```
SELECT t.cog, DATE_FORMAT(t.o_date, '%y%m') as YearMonth, SUM(t.price) as sales
FROM
(SELECT o.*, t.cog
FROM main o
JOIN user_cog t
ON o.user_id = t.user_id) t
GROUP BY t.cog, DATE_FORMAT(t.o_date, '%y%m');
```

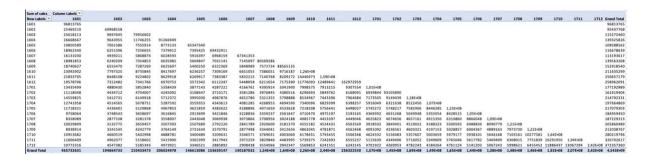
3. Получим объем спроса каждой новой когорты в 1 месяц существования, для прогнозирования спроса новый когорты 1801

```
SELECT t.cog, DATE_FORMAT(t.o_date, '%y%m') as YearMonth, SUM(t.price) as sales
FROM
(SELECT o.*, t.cog
FROM main o
JOIN user_cog t
ON o.user_id = t.user_id) t
GROUP BY t.cog, DATE_FORMAT(t.o_date, '%y%m')
HAVING cog=YearMonth;
```

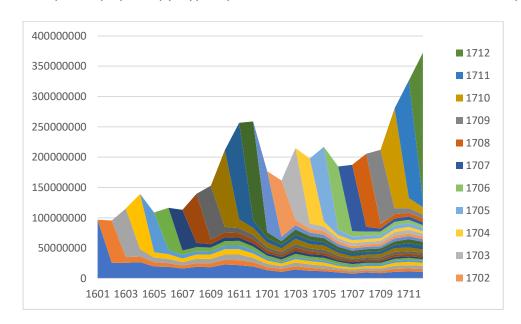
- 4. Выгрузим полученные данные в Excel
- 5. Спрогнозируем объем спроса новой когорты 1801 в период 1801 с учетом сезонности.

				1. Рассчитываем зна	чения функции у=bx+a							
				b	a							
				4360226,39	60338909,37							
Продаж	и когорт в перв	ый месяц		2. Значения тренда	3. Отклонение фактических значений от значений тренда	4. Среднее отклонение для каждого месяца	5. Общий индекс сезонности	для какого месяца коэффициент	6. Коэффициенты сезонности очищенные от роста	7. Задаем период, на который хотим рассчитать прогноз	8. Рассчитыва ем значения тренда для 1801	9. Расчет прогноза
2016		96813765	1	64699135,76	1,496368751		1,006025	1	1,171293201	25	1,69E+08	198 352 142
2016	2	69968558	2	69059362,15	1,013165425	0,890107		2	0,884776345			
2016	3	79956602	3	73419588,54	1,089036367	1,013476		3	1,007406368			
2016	4	91266949	4	77779814,93	1,173401468	0,998554		4	0,992574215			
2016	5	65347340	5	82140041,32	0,795560106	0,85565		5	0,850525707			
2016	6	69432911	6	86500267,71	0,802690128	0,779449		6	0,774781245			
2016	7	67341353	7	90860494,1	0,741151076	0,751195		7	0,746696408			
2016	8	80569286	8	95220720,49	0,846131867	0,831924		8	0,826942076			
2016	9	88565135	9	99580946,88				9	0,840158234			
2016			10	103941173,3				10				
2016	11	159238214	11	108301399,7	1,470324617	1,339736		11	1,331712695			
2016	12	162972959	12	112661626,1	1,446570272	1,456297		12	1,447575565			
2017	1	100677498	13	117021852,4	0,860330749							
2017	2	93105890	14	121382078,8	0,767048076							
2017	3	117935565	15	125742305,2	0,937914768							
2017	4	107166320	16	130102531,6	0,823706646							
2017	5	123132849	17	134462758	0,915739427							
2017	6	104979027	18	138822984,4	0,756207824							
2017	7	108996616	19	143183210,8	0,761238803							
2017	8	120648665	20	147543437,2	0,817716242							
2017	9	121684154	21	151903663,6	0,801061351							
2017	10	164020311	22	156263890	1,049636682							
2017	11	194218135	23	160624116,3	1,209146792							
2017	12	241870829	24	164984342,7	1,466022927							

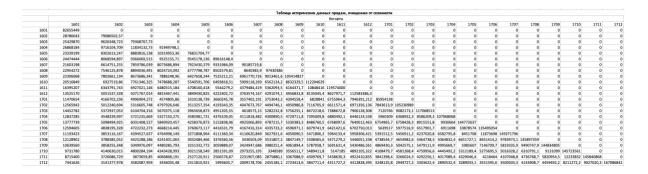
6. Построим сводную таблицу по полученным из SQL данным (Строки – период времени, колонки – когорты)



7. Построим график структуры продаж компании за 2016-2017 год по когортам.



8. Очистим исторические данные от сезонности, чтобы построить более точный прогноз на период 1801.



9. Определим коэффициенты выживаемости когорт по очищенным от сезонности данным, найдем средневзвешенный коэффициент выживаемости на каждый период.

Териод	Среда	невзв. Коэф.									Ko	эф. Выживая	емости, очище	нный от сезс	онности											
	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
	2 0	0,100691401	0,348265524	0,121728472	0,149103143	0,112180333	0,12423489	0,104132841	0,0958066	0,0925118	0,0818891	0,0966842	0,072155918	0,0705823	0,0912535	0,086307	0,0733435	0,0919778	0,0751425	0,087632	0,0670265	0,0650509	0,0630449	0,0839523	0,0618964	
	3 0	0,082815966	0,307661146	0,122863467	0,111930882	0,103590884	0,10312709	0,094877764	0,0771937	0,0673519	0,0584773	0,0725019	0,056500278	0,0619217	0,0828438	0,0605234	0,0761567	0,0640161	0,0583792	0,0623743	0,0398662	0,043255	0,0401903	0,0563483		
	4 0	0,076596916	0,325062471	0,104989384	0,118006992	0,087805427	0,08821617	0,083971609	0,0655217	0,0550351	0,0470222	0,0608606	0,055292718	0,0702246	0,0693429	0,0642247	0,055676	0,0609698	0,045946	0,0527452	0,0343651	0,0324664	0,0307902			
	5 0	0,070756579	0,282367342	0,102030142	0,099028702	0,087272932	0,08365893	0,072462593	0,0485607	0,0440425	0,0408345	0,0634261	0,061791633	0,0572564	0,078588	0,0568038	0,0532767	0,0469105	0,0413449	0,0389364	0,0281359	0,0284029				
	6 0	0,070626696	0,296102001	0,083645792	0,112059652	0,085766945	0,07211335	0,061866966	0,0417314	0,0385203	0,0469667	0,0755069	0,057541707	0,0678609	0,0656108	0,0518565	0,0466798	0,0462699	0,0366933	0,0311337	0,0239807					
	7 0	0,068282309	0,26136447	0,095423342	0,109283369	0,081334516	0,06127887	0,047116539	0,0394453	0,0461329	0,0546057	0,0659065	0,05660175	0,0595359	0,0627255	0,0442359	0,0468026	0,0411741	0,0299024	0,026212						
	8 0	0,068555756	0,277710326	0,098667319	0,097788419	0,070476704	0,05010211	0,040843596	0,0448815	0,0542155	0,0448676	0,0756365	0,05370249	0,0529235	0,0551288	0,0479191	0,0406596	0,0372063	0,0227186							
	9 0	0,065560723	0,269868094	0,087600858	0,08727641	0,052370643	0,04308298	0,046792274	0,0512118	0,0441432	0,0505658	0,0541844	0,049878179	0,0483735	0,0566515	0,0435354	0,0366645	0,0267721								
1	0 0	0,062081319	0,24822137	0,080219416	0,063901295	0,049753327	0,04572666	0,055249033	0,0455928	0,0491347	0,0436084	0,0487663	0,045756429	0,0482352	0,0477985	0,0310367	0,0282196									
1	1 0	0,059158439	0,198356032	0,063878417	0,062958068	0,052179878	0,05161488	0,048831702	0,0546103	0,0445007	0,0384205	0,049093	0,047599239	0,0434532	0,0446913	0,028031										
1	2 0	0,057324859	0,163633136	0,052689375	0,066986119	0,061668196	0,0466732	0,057499001	0,0462093	0,037236	0,0364953	0,0515759	0,043046022	0,0437224	0,0377892											
1	3 0	0,057904952	0,138776743	0,063381498	0,0769157	0,057935453	0,05542473	0,049282767	0,045646	0,0362495	0,0385325	0,0489232	0,045483385	0,0383079												
1	4 0	0,057960579	0,15202799	0,067942627	0,072084219	0,063522963	0,04894689	0,04587745	0,0381047	0,0408523	0,0337382	0,0384099	0,03605907													
1	5 0	0,057797352	0,174855773	0,057515315	0,075992216	0,050769192	0,0425841	0,037678037	0,0379766	0,034369	0,0319413	0,0342919														
1	6 0	0,054781816	0,155310813	0,064472212	0,059497362	0,051821764	0,03946918	0,034144394	0,0329681	0,0295173	0,0258352															
1	7 0	0,056681279	0,166688815	0,061180635	0,055368475	0,046127381	0,04231759	0,031820059	0,0247479	0,0251994																
1	8	0,05654569	0,152374774	0,048723972	0,05575604	0,048725373	0,03932177	0,028639669	0,0222782																	
1	9 0	0,058221586	0,135422692	0,047905425	0,063626742	0,047236986	0,03289164	0,022246032																		
2	0 0	0,066065751	0,146740496	0,048788958	0,06048077	0,044229248	0,03008928																			
2	1 0	0,066813994	0,128721824	0,052359683	0,048805846	0,037368623																				
2	2 0	0,069996324	0,11773912	0,047117641	0,045132211																					
2	3 0	0,072583168	0,105442536	0,039723799																						
2	4 0	0,090031887	0,090031887																							

10. Построим коэффициент выживаемости когорты на 25 период (Используем только значения за последние несколько лет для более точного прогноза, так как коэффициент выживаемости уменьшается нелинейно)

Прогноз коэф.	виживаемости на	25 год для когорты 1601	
		b	a
21	0,128721824	-0,01283	0,399308225
22	0,11773912		
23	0,105442536		
24	0,090031887		
25	0,078392244		

11. Построим прогноз по выручке на период 1801 с учетом сезонности.

Cohort	Sales 1st period	Sales no seasonality	Seasonality rate	Sales	Period for pred.	Survival rate	Prediction
1601	96813765	82655448,62	1,171293201	96813765	25	0,078392244	758944
1602	69968558	79080502,57	1,171293201	92626455	24	0,090031887	833933
1603	79956602	79368767,73	1,171293201	92964098	23	0,072583168	674762
1604	91266949	91949748,1	1,171293201	107700115	22	0,069996324	753861
1605	65347340	76831704,77	1,171293201	89992453	21	0,066813994	601275
1606	69432911	89616148,36	1,171293201	104966785	20	0,066065751	693471
1607	67341353	90185719,8	1,171293201	105633920	19	0,058221586	61501
1608	80569286	97430386,4	1,171293201	114119549	18	0,05654569	645296
1609	88565135	105414827,1	1,171293201	123471670	17	0,056681279	69985
1610	126292811	112204628,7	1,171293201	131424519	16	0,054781816	71996
1611	159238214	119574000,1	1,171293201	140056213	15	0,057797352	80948
1612	162972959	112583386,3	1,171293201	131868155	14	0,057960579	76431
1701	100677498	85954138,47	1,171293201	100677498	13	0,057904952	58297
1702	93105890	105230989,2	1,171293201	123256342	12	0,057324859	70656
1703	117935565	117068512,6	1,171293201	137121553	11	0,059158439	81118
1704	107166320	107968067,7	1,171293201	126462264	10	0,062081319	78509
1705	123132849	144772636,5	1,171293201	169571205	9	0,065560723	111172
1706	104979027	135495054,5	1,171293201	158704436	8	0,068555756	108801
1707	108996616	145971796,4	1,171293201	170975773	7	0,068282309	116746
1708	120648665	145897359,1	1,171293201	170888585	6	0,070626696	120692
1709	121684154	144834805	1,171293201	169644022	5	0,070756579	120034
1710	164020311	145723560,6	1,171293201	170685016	4	0,076596916	130739
1711	194218135	145840867,7	1,171293201	170822417	3	0,082815966	141468
1712	241870829	167086841,5	1,171293201	195707681	2	0,100691401	197060
1801	198352142				1	1	1983521
					Total sale	es 1801	4175837
					Old	d	2192316
					Nev	N	1983521

12. Построим график продаж по историческим данным и прогнозируемому значению на период 1801.

