

Документ подписан простой электронной подписью

Информация о владельце:

ФИО: Локтионова Оксана Геннадьевна

Должность: проректор по учебной работе

Дата подписания: 08.10.2023 16:33:43

Уникальный идентификатор:

0b817ca911e6668abb13a5d426d39e5f1c11eabbf73e943df4a4851fda56d089


МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное образовательное учреждение высшего образования

«Юго-Западный государственный университет»

Кафедра «Машиностроительные технологии и оборудование»

УТВЕРЖДАЮ:
Проректор по учебной работе
О.И. Локтионова
« 4 » 02 20



КОРРЕЛЯЦИОННЫЙ АНАЛИЗ СВЯЗЕЙ МЕЖДУ КОЛИЧЕСТВЕННЫМИ ПРИЗНАКАМИ

Методические указания к проведению практических занятий для студентов по направлению подготовки 15.03.05 Конструкторско-технологическое обеспечение машиностроительных производств профиль «Технология машиностроения»

Курск 2018

УДК 519.6

Составители: В.В. Куц, М.С. Разумов

Рецензент

Кандидат технических наук, доцент *А.О. Гладышкин*

Корреляционный анализ связей между количественными признаками : методические указания к проведению практических и лабораторных занятий / Юго-Зап. гос. ун-т; сост.: В.В. Куц, М.С. Разумов. – Курск, 2018. 12 с.: ил. 2.: табл. 2.

Содержат сведения по вопросам выполнения корреляционного анализа связей между количественными признаками. Указывается порядок выполнения практического занятия, подходы к решению и правила оформления.

Методические рекомендации соответствуют требованиям программы, утвержденной учебно-методическим объединением по специальности автоматизированного машиностроительного производства (УМОАМ).

Предназначено для студентов направлений 15.04.05 «Конструкторско-технологическое обеспечение машиностроительных производств» профиль «Технология машиностроения» дневной и заочной форм обучения.

Текст печатается в авторской редакции

Подписано в печать 07.02.18 г. Формат 60x84 1/16.
Усл.печ.л. 0,7. Уч.-изд.л. 0,5. Тираж 40 экз. Заказ. 75 Бесплатно.
Юго-Западный государственный университет.
305040, г. Курск, ул. 50 лет Октября, 94.

1 Цель работы: исследовать зависимости между результатами наблюдений и определить их характер.

2 Задание: построить диаграммы рассеивания, рассчитать коэффициенты корреляции для результатов наблюдения трёх случайных величин, а установить значимость этих коэффициентов.

3 Краткие теоретические сведения

В технологической практике исследователь часто сталкивается с необходимостью установления факта существования функциональных или иных зависимостей между экспериментальными данными, нередко такая связь может быть случайной. Различные постановки задач статистического исследования можно классифицировать следующим образом: задачи корреляционного анализа (задачи исследования наличия взаимосвязей между отдельными группами переменных); задачи регрессионного анализа; задачи дисперсионного анализа.

3.1 Понятие о корреляционном анализе

Корреляция (от лат. *Correlatio* – соотношение) – мера связи. Практическая статистика использует это понятие для обозначения группы показателей, описывающих степень сходства в варьировании двух и более переменных. Предварительный анализ зависимости между оценками параметра дают простые коэффициенты корреляций. Коэффициент корреляции является мерой линейной зависимости между анализируемыми переменными. Кроме того, в статистическом анализе используются коэффициенты частной корреляции, которые исключают зависимость остальных переменных, кроме рассматриваемой. Все эти корреляции весьма полезны для первичного анализа взаимозависимости. Корреляционный анализ определяет меру линейности зависимости независимо от положения прямой, то есть наличие некоторой гипотетической функциональной зависимости.

Оценкой коэффициента корреляции двух случайных величин X и Y является величина:

$$r_{XY} = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2 \sum_{i=1}^n (Y_i - \bar{Y})^2}},$$

где n - объем выборки X и Y ;

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i;$$

$$\bar{Y} = \frac{1}{n} \sum_{i=1}^n Y_i$$

Сильная корреляция означает закономерность: чем более отличается от своего среднего значение одной переменной, тем дальше от своего среднего значение другой. Также можно утверждать, что чем ближе точки на графике рассеивания параметров X и Y расположены к прямой линии больше корреляция между X и Y .

Коэффициент может принимать как положительные, так и отрицательные значения. Знак означает направленность изменчивости двух переменных. Если коэффициент корреляции больше нуля, то для данного наблюдения отличия от среднего бывают, как правило, в одну сторону. Если меньше нуля – увеличение первой переменной, как правило, связано с уменьшением второй. Сравнение именно со своими средними позволяет сравнивать переменные разной размерности и природы, например метры с килограммами или секунды с часами. Нас может интересовать, в зависимости от цели исследования, отличие коэффициента корреляции от нуля (есть ли вообще связь между переменными) или отличие двух коэффициентов друг от друга (одинаковы ли связи, например, между переменными A и B , с одной стороны и переменными C и D , с другой).

Нельзя сразу принимать полученный коэффициент корреляции как окончательный результат, если исследование отвечает на сложные вопросы. Если вероятность нуль-гипотезы достаточно мала, то можно сказать только, что между переменными существует какая-то связь, и не более того. Ответы на другие вопросы – предмет дальнейшего анализа, а часто – и специальных опытов.

Коэффициент корреляции находится в пределе $|r| \leq 1$. Абсолютная величина коэффициента означает силу связи (условно принима-

ем: 0 – 0,3 – слабая связь; 0,3 – 0,7 – связь средней силы; 0,7 – 1 – сильная связь).

Знак коэффициента определяет, совпадают ли направления изменчивости переменных. «+» – совпадают (чем больше одно, тем больше другое), «-» – противоположны (чем больше одно, тем меньше другое). На рис. 1. приведены различные случаи корреляционной связи.

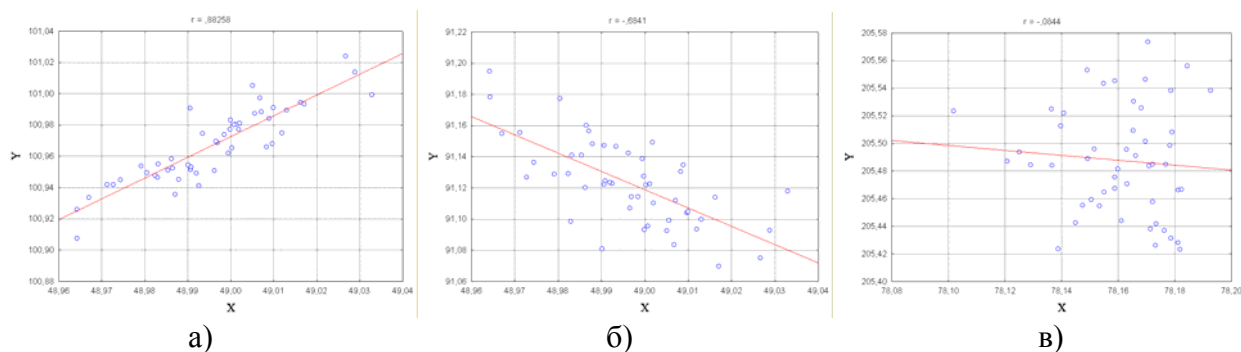


Рис. 1 – Графики рассеивания с различной силой корреляционной связи:
а – сильная, положительная; б – средняя отрицательная;
в – слабая отрицательная

3.2 Проверка значимости коэффициентов корреляции

Для проверки значимости коэффициентов корреляции чаще всего используют распределение Стьюдента и условие

$$\frac{|r_{XY}| \sqrt{n-2}}{\sqrt{1-r_{XY}^2}} < t_{P,k},$$

где $t_{P,k}$ – коэффициент Стьюдента, который берется из справочных таблиц (приложение А) при заданной доверительной вероятности $P=1-q$ и числе степеней свободы $k=n-2$.

Если условие выполняется, то гипотеза об отсутствии корреляционной связи принимается.

3.3. Пример выполнения работы

В таблице представлены результаты 15-ти наблюдений двух случайных величин X и Y .

	Номер наблюдения														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
X	5,48	5,57	5,49	5,40	5,55	5,41	5,54	5,46	5,33	5,41	5,55	5,53	5,42	5,60	5,35
Y	1,61	1,54	1,56	1,47	1,60	1,51	1,61	1,45	1,43	1,43	1,60	1,56	1,47	1,61	1,48

Для расчета коэффициента корреляции заполним следующую таблицу

№	X_i	$X_i - \bar{X}$	$(X_i - \bar{X})^2$	Y_i	$Y_i - \bar{Y}$	$(Y_i - \bar{Y})^2$	$(X_i - \bar{X})(Y_i - \bar{Y})$
1	5,48	0,01	0,0001	1,61	0,08	0,0066	0,0006
2	5,57	0,10	0,0095	1,54	0,01	0,0001	0,0011
3	5,49	0,02	0,0003	1,56	0,03	0,0010	0,0005
4	5,40	-0,07	0,0053	1,47	-0,06	0,0034	0,0043
5	5,55	0,08	0,0060	1,60	0,07	0,0051	0,0055
6	5,41	-0,06	0,0039	1,51	-0,02	0,0003	0,0012
7	5,54	0,07	0,0045	1,61	0,08	0,0066	0,0055
8	5,46	-0,01	0,0002	1,45	-0,08	0,0062	0,0010
9	5,33	-0,14	0,0204	1,43	-0,10	0,0097	0,0141
10	5,41	-0,06	0,0039	1,43	-0,10	0,0097	0,0062
11	5,55	0,08	0,0060	1,60	0,07	0,0051	0,0055
12	5,53	0,06	0,0033	1,56	0,03	0,0010	0,0018
13	5,42	-0,05	0,0028	1,47	-0,06	0,0034	0,0031
14	5,60	0,13	0,0162	1,61	0,08	0,0066	0,0104
15	5,35	-0,12	0,0150	1,48	-0,05	0,0024	0,0060
	$\Sigma=82,09$		$\Sigma=0,0973$	$\Sigma=22,93$		$\Sigma=0,0674$	$\Sigma=0,0667$
	$\bar{X} = 5,47$			$\bar{Y} = 1,53$			

Рассчитаем величину коэффициента корреляции

$$r_{XY} = \frac{0,0667}{\sqrt{0,0973 \cdot 0,0674}} = 0,83.$$

На рис. 2 показана диаграмма рассеивания результатов наблюдения X и Y .

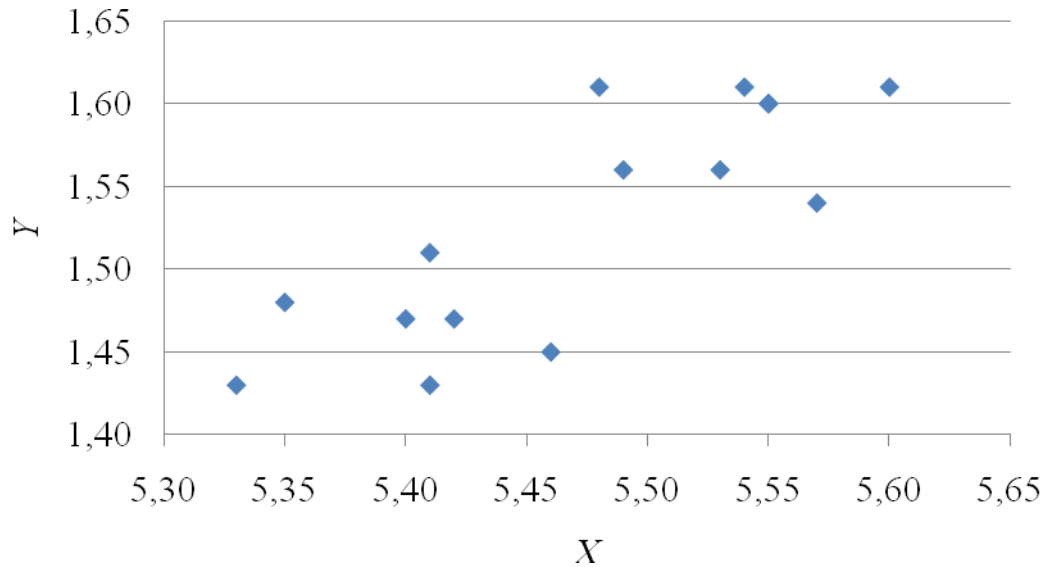


Рис. 2 диаграмма рассеивания результатов наблюдения X и Y

Исходя из рассчитанного коэффициента корреляции и диграммы рассеивания, можно сделать вывод, что между наблюдениями X и Y существует сильная положительная связь.

Далее выполним проверку значимости коэффициентов корреляции при уровне значимости $q=0,05$

$$t = \frac{|r_{XY}| \sqrt{n-2}}{\sqrt{1-r_{XY}^2}} = \frac{|0,83| \sqrt{15-2}}{\sqrt{1-0,83^2}} = 5,23 > t_{0,95;13} = 2,160,$$

где $t_{0,95;13}$ – коэффициент Стьюдента из справочной таблицы приложения А, взятый при заданной доверительной вероятности $P=1-q=1-0,05=0,95$ и числе степеней свободы $k=n-2=15-2=13$.

Т.к. условие выполняется, то можно сделать вывод о значимости коэффициента корреляции между наблюдениями X и Y .

4 Выполнение работы

Получив исходные данные для выполнения практической работы (см. приложение Б), студент изучает теоретические сведения согласно пункту 3. Далее выполняет расчеты аналогичные в рассмотренном примере с учетом имеющихся особенностей задания.

В отчёте по практической работе должны найти отражение следующие пункты:

- название работы;
- цель работы;
- индивидуальное задание для выполнения работы;
- краткие теоретические сведения;
- результаты выполнения работы;
- подробные выводы по работе.

Контрольные вопросы

1. Что такое корреляция?
2. Что такое коэффициент корреляции?
3. Как рассчитывается коэффициент корреляции?
4. Какова область значений коэффициента корреляции?
5. Что означает положительная и отрицательная корреляция?
6. При каком значении коэффициента корреляции связь является сильной, средней и слабой?

Библиографический список

1. Сергеев, А.Г. Метрология [Текст]/ А.Г. Сергеев, В.В. Крохин. Учебное пособие для вузов. М.: Логос, 2001. 488 с.: ил.
2. Алексахин, С.В. Прикладной статистический анализ [Текст]/ С.В. Алексахин, А.В. Балдин, А.Б. Николаев, В.Ю. Строганов. Учебное пособие для вузов. М.: “Издательство ПРИОР”, 2001. 224 с.

ПРИЛОЖЕНИЕ А
Распределение Стьюдента (t_p)

k	P											
	0,05	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	0,95	0,99
1	0,079	0,158	0,325	0,510	0,727	1,000	1,376	1,963	3,078	6,314	12,706	63,657
2	0,071	0,142	0,289	0,445	0,617	0,816	1,061	1,386	1,886	2,920	4,303	9,925
3	0,068	0,137	0,277	0,424	0,584	0,765	0,978	1,250	1,638	2,353	3,182	5,841
4	0,067	0,134	0,271	0,414	0,569	0,741	0,941	1,190	1,533	2,132	2,776	4,604
5	0,066	0,132	0,267	0,408	0,559	0,727	0,920	1,156	1,476	2,015	2,571	4,032
6	0,065	0,131	0,265	0,404	0,553	0,718	0,906	1,134	1,440	1,943	2,447	3,707
7	0,065	0,130	0,263	0,402	0,549	0,711	0,896	1,119	1,415	1,895	2,365	3,499
8	0,065	0,130	0,262	0,399	0,546	0,706	0,889	1,108	1,397	1,860	2,306	3,355
9	0,064	0,129	0,261	0,398	0,543	0,703	0,883	1,100	1,383	1,833	2,262	3,250
10	0,064	0,129	0,260	0,397	0,542	0,700	0,879	1,093	1,372	1,812	2,228	3,169
11	0,064	0,129	0,260	0,396	0,540	0,697	0,876	1,088	1,363	1,796	2,201	3,106
12	0,064	0,128	0,259	0,395	0,539	0,695	0,873	1,083	1,356	1,782	2,179	3,055
13	0,064	0,128	0,259	0,394	0,538	0,694	0,870	1,079	1,350	1,771	2,160	3,012
14	0,064	0,128	0,258	0,393	0,537	0,692	0,868	1,076	1,345	1,761	2,145	2,977
15	0,064	0,128	0,258	0,393	0,536	0,691	0,866	1,074	1,341	1,753	2,131	2,947
16	0,064	0,128	0,258	0,392	0,535	0,690	0,865	1,071	1,337	1,746	2,120	2,921
17	0,064	0,128	0,257	0,392	0,534	0,689	0,863	1,069	1,333	1,740	2,110	2,898
18	0,064	0,127	0,257	0,392	0,534	0,688	0,862	1,067	1,330	1,734	2,101	2,878
19	0,064	0,127	0,257	0,391	0,533	0,688	0,861	1,066	1,328	1,729	2,093	2,861
20	0,063	0,127	0,257	0,391	0,533	0,687	0,860	1,064	1,325	1,725	2,086	2,845
21	0,063	0,127	0,257	0,391	0,532	0,686	0,859	1,063	1,323	1,721	2,080	2,831
22	0,063	0,127	0,256	0,390	0,532	0,686	0,858	1,061	1,321	1,717	2,074	2,819
23	0,063	0,127	0,256	0,390	0,532	0,685	0,858	1,060	1,319	1,714	2,069	2,807
24	0,063	0,127	0,256	0,390	0,531	0,685	0,857	1,059	1,318	1,711	2,064	2,797
25	0,063	0,127	0,256	0,390	0,531	0,684	0,856	1,058	1,316	1,708	2,060	2,787
26	0,063	0,127	0,256	0,390	0,531	0,684	0,856	1,058	1,315	1,706	2,056	2,779
27	0,063	0,127	0,256	0,389	0,531	0,684	0,855	1,057	1,314	1,703	2,052	2,771
28	0,063	0,127	0,256	0,389	0,530	0,683	0,855	1,056	1,313	1,701	2,048	2,763
29	0,063	0,127	0,256	0,389	0,530	0,683	0,854	1,055	1,311	1,699	2,045	2,756
30	0,063	0,127	0,256	0,389	0,530	0,683	0,854	1,055	1,310	1,697	2,042	2,750
300	0,063	0,126	0,254	0,386	0,525	0,675	0,843	1,038	1,284	1,650	1,968	2,592

ПРИЛОЖЕНИЕ Б
Исходные данные к работе

Таблица 1

	№ варианта														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
<i>q</i>	0,05	0,1	0,05	0,1	0,05	0,1	0,05	0,1	0,05	0,1	0,05	0,1	0,05	0,1	0,05

Таблица 2

№	№ варианта								
	1			2			3		
1	178,435	202,092	144,037	194,150	14,380	12,011	113,620	138,372	10,443
2	178,416	201,730	143,846	194,104	14,325	12,009	113,616	138,355	10,668
3	178,370	201,814	143,907	194,116	14,357	12,005	113,614	138,368	10,617
4	178,457	201,802	143,892	194,113	14,317	12,013	113,606	138,362	10,622
5	178,398	201,720	143,924	194,122	14,345	11,987	113,624	138,374	10,685
6	178,444	201,868	143,899	194,116	14,343	12,002	113,585	138,346	10,463
7	178,351	202,009	144,018	194,142	14,354	11,978	113,610	138,371	10,485
8	178,465	201,810	143,978	194,135	14,352	11,994	113,631	138,358	10,483
9	178,466	201,773	143,980	194,129	14,394	12,000	113,641	138,381	10,554
10	178,490	201,816	144,029	194,144	14,375	11,991	113,628	138,364	10,458
11	178,442	201,754	143,880	194,111	14,306	11,995	113,634	138,375	10,465
12	178,510	202,089	144,062	194,153	14,426	11,969	113,707	138,418	10,637
13	178,512	201,997	144,099	194,162	14,408	11,994	113,621	138,376	10,469
14	178,360	201,826	144,041	194,144	14,391	11,984	113,619	138,370	10,428
15	178,417	201,812	144,059	194,155	14,393	11,991	113,650	138,380	10,539

№	№ варианта								
	4			5			6		
1	49,293	142,859	169,932	197,329	23,619	203,924	61,057	34,567	206,450
2	49,353	142,918	169,724	197,099	23,735	203,934	61,220	34,554	206,483
3	49,336	142,788	169,858	197,098	23,761	203,950	61,146	34,588	206,380
4	49,333	142,860	169,732	197,095	23,826	203,944	61,149	34,564	206,372
5	49,340	142,836	169,747	197,119	23,690	203,923	61,269	34,629	206,326
6	49,318	142,894	169,765	197,029	23,763	203,929	61,147	34,588	206,444
7	49,324	142,829	169,923	197,097	23,702	203,928	61,088	34,592	206,411
8	49,316	142,862	169,766	197,196	23,729	203,935	61,235	34,584	206,376
9	49,325	142,843	169,602	196,985	23,723	203,920	61,233	34,602	206,288
10	49,317	142,816	169,630	197,146	23,699	203,938	61,280	34,565	206,391
11	49,321	142,881	169,856	197,242	23,698	203,925	61,152	34,575	206,378
12	49,338	142,897	169,697	197,101	23,714	203,923	61,189	34,544	206,465
13	49,323	142,934	169,603	197,085	23,725	203,927	61,200	34,602	206,391

№ варианта									
№	4			5			6		
14	49,322	142,859	169,691	197,210	23,677	203,930	61,238	34,591	206,372
15	49,322	142,928	169,860	197,128	23,681	203,931	61,241	34,632	206,318

№ варианта									
№	7			8			9		
1	55,995	80,812	206,307	138,182	183,767	44,262	145,114	154,467	191,922
2	56,028	80,817	206,274	138,133	183,792	44,302	145,091	154,464	191,933
3	55,948	80,732	206,265	138,204	183,777	44,260	145,104	154,479	191,945
4	55,986	80,935	206,242	138,148	183,797	44,263	145,113	154,441	191,904
5	56,096	80,891	206,266	138,156	183,806	44,291	145,104	154,429	191,921
6	56,098	80,818	206,255	138,068	183,798	44,280	145,106	154,425	191,908
7	55,984	80,738	206,281	138,194	183,765	44,257	145,113	154,375	191,910
8	55,846	80,690	206,254	138,217	183,775	44,282	145,104	154,423	191,906
9	56,117	80,934	206,316	138,155	183,788	44,294	145,083	154,365	191,887
10	56,057	80,826	206,266	138,129	183,785	44,274	145,118	154,444	191,891
11	56,062	80,776	206,274	138,165	183,776	44,227	145,127	154,468	191,919
12	55,858	80,735	206,264	138,091	183,814	44,208	145,163	154,463	191,929
13	56,076	80,752	206,283	138,141	183,803	44,235	145,136	154,464	191,929
14	56,025	80,861	206,281	138,170	183,798	44,261	145,103	154,415	191,913
15	55,851	80,832	206,248	138,189	183,771	44,246	145,129	154,339	191,901

№ варианта									
№	10			11			12		
1	198,209	43,903	200,697	152,883	155,297	114,758	50,691	195,826	96,043
2	198,198	43,910	200,691	152,882	155,270	114,765	50,612	195,813	95,961
3	198,191	43,937	200,693	152,868	155,253	114,762	50,498	195,780	95,920
4	198,212	43,910	200,697	152,882	155,172	114,743	50,751	195,837	96,074
5	198,203	43,933	200,694	152,912	155,317	114,742	50,615	195,809	95,988
6	198,212	43,902	200,686	152,840	155,303	114,760	50,651	195,818	96,005
7	198,210	43,902	200,696	152,870	155,177	114,750	50,618	195,807	95,969
8	198,211	43,936	200,683	152,889	155,339	114,759	50,577	195,799	95,934
9	198,227	43,898	200,692	152,898	155,270	114,763	50,613	195,809	95,986
10	198,221	43,911	200,683	152,867	155,273	114,761	50,609	195,802	95,958
11	198,208	43,922	200,684	152,876	155,272	114,744	50,637	195,817	95,997
12	198,200	43,898	200,696	152,888	155,270	114,766	50,527	195,795	95,914
13	198,200	43,920	200,691	152,894	155,340	114,750	50,643	195,817	95,998
14	198,209	43,910	200,686	152,874	155,337	114,746	50,524	195,788	95,925
15	198,218	43,875	200,688	152,898	155,359	114,746	50,576	195,791	95,964

№ варианта									
№	13			14			15		
1	206,732	118,321	182,070	190,116	180,982	83,113	117,240	114,728	139,641
2	206,735	118,275	181,949	190,122	181,088	83,153	117,263	114,733	139,626
3	206,773	118,307	182,053	190,117	180,977	83,117	117,242	114,730	139,637
4	206,675	118,325	182,000	190,117	180,978	83,120	117,249	114,738	139,641
5	206,761	118,298	182,005	190,121	181,039	83,145	117,258	114,735	139,636
6	206,727	118,302	181,996	190,129	181,229	83,223	117,180	114,731	139,636
7	206,748	118,315	181,988	190,118	181,054	83,124	117,281	114,738	139,624
8	206,639	118,283	181,984	190,117	181,039	83,138	117,203	114,728	139,645
9	206,622	118,313	182,020	190,117	181,014	83,136	117,330	114,734	139,649
10	206,688	118,298	181,996	190,114	180,985	83,127	117,321	114,733	139,635
11	206,690	118,290	181,955	190,117	181,042	83,173	117,181	114,731	139,631
12	206,661	118,290	181,968	190,123	181,085	83,161	117,306	114,728	139,649
13	206,749	118,297	181,969	190,119	181,041	83,132	117,329	114,724	139,649
14	206,717	118,318	182,034	190,121	181,038	83,151	117,232	114,732	139,633
15	206,695	118,281	181,924	190,118	181,046	83,150	117,246	114,728	139,643