



TQU VERLAG

**Grafiken
für multivariate
Datensätze**

Multivariate Grafiken

Qualität verstehen durch Berücksichtigung von Korrelationen

[Dr. Konrad Reuter](#)

Für die Darstellung Multivariater Datensätze haben sich eine Reihe von grafischen Methoden etabliert, die in der Regel von Statistiksoftware angeboten werden.

Hier sollen für EXCEL-analoge Grafiken herigestellt werden.

Boxplots dienen vorrangig dem Vergleich univariater Variablen. Die Darstellung vermittelt übersichtlich die wichtigsten Kennwerte von Variablen. Statistische Maßzahlen im Vergleich, insbesondere in zeitlichen Entwicklungen, suchen Veränderungen, die Datensätze für Lösungsansätze von Sachverhaltsproblemen bieten können. Scatterplots bieten auf einen Blick eine Übersicht zu Korrelationen in multivariaten Datensätzen.

Jeder Anwendungsfall muss individuell behandelt werden. Der Autor hat, basierend auf speziellen Fällen aus seiner Praxis, in diesem QUALITY APP eine praktische und sehr anschauliche Vorgehensweise für Multivariate Darstellungen entwickelt, die eine möglichst vollständige Darstellung vernachlässigen. Mit diesem QUALITY APP können interessante mehrdimensionale Variable in ihrer Wirkung beobachtet und verstanden werden. Es unterstützt Personen, die mit multivariaten Datensätzen umgehen müssen oder sich mit den Ursachen für nicht einfach zu klärende Herstell- und Funktionsprobleme beschäftigen müssen.

Das APP ist in Excel programmiert und kann sofort eingesetzt werden.

TQU Verlag, Magirus-Deutz-Straße 18, 89077 Ulm Deutschland, Telefon 0731/14660200, verlag@tqu-group.com, www.tqu-verlag.com

QUALITY APPs Applikationen für das Qualitätsmanagement

Lizenzvereinbarung

Dieses Produkt "Multivariate Grafiken" wurde von uns mit großem Aufwand und großer Sorgfalt hergestellt. Dieses Werk ist urheberrechtlich geschützt (©). Die dadurch begründeten Rechte, insbesondere die der Weitergabe, der Übersetzung, des Kopierens, der Entnahme von Teilen oder der Speicherung bleiben vorbehalten.

Bei Fehlern, die zu einer wesentlichen Beeinträchtigung der Nutzung dieses Softwareproduktes führen, leisten wir kostenlos Ersatz. Beschreibungen und Funktionen verstehen sich als Bereinigung von Nutzer-Smögkkeiten und nicht als rechtlich verbindliche Nutzungsbestimmte Eigenschaften. Wir übernehmen keine Gewähr dafür, dass die angebotenen Lösungen für bestimmte Voraussetzungen bei bestimmten beabsichtigten Zwecken geeignet sind.

Sie erklären sich damit einverstanden, dieses Produkt nur für Ihre eigene Arbeit und für die Information innerhalb Ihres Unternehmens zu verwenden. Sollten Sie es in anderer Form, insbesondere in Schulungs- und Informationsmaßnahmen bei anderen Unternehmen (Beratung, Schulungseinrichtung etc.) verwenden wollen, setzen Sie sich unbedingt vorher mit uns in Verbindung. Das Produkt wird von uns regelmäßig weiterentwickelt. Bitte melden Sie sich, wenn Sie ein Update wünschen.

Alle Ergebnisse basieren auf den vom Autor eingesetzten Formeln und müssen vom Anwender sorgfältig geprüft werden. Die berechneten Ergebnisse sind als Hinweise und Anregungen zu verstehen.

Wir wünschen viel Spaß und Anregungen mit dieser Applikation

TQU Verlag, Magirus-Deutz-Straße 18, 89077 Ulm Deutschland, Telefon 0731/14660200, verlag@tqu-group.com, www.tqu-verlag.com

Grafiken für multivariate Datensätze

1. Statistik

Blatt

1.1 Motivation

Qualität bedeutet die simultane Erfüllung aller Kundenanforderungen.

Die gewöhnlich erfassten Merkmale von Produkten sind in der Regel nicht unabhängig voneinander.

Die Missachtung dieser Abhängigkeiten führt nicht selten zu Fehlentscheidungen, die mit erheblichen Kosten verbunden sein können.

Das Erkennen von Abhängigkeiten in Datenstrukturen ist für die weitere realistischere Berechnung zu Toleranzen, Prozessfähigkeiten, Regelkarten und der Prozesssteuerung unerlässlich.

Diese Abhängigkeitsstrukturen sind in der Regel sehr stabil, da sie durch die naturgesetzliche Ursache/Wirkungsbeziehung geprägt sind.

Zum Verständnis von Abhängigkeiten dienen eine Reihe von Grafiken, die hier vorgestellt werden sollen.

Die volle Flexibilität bieten Statistikprogramme, deren Leistungsfähigkeit hier mit EXCEL nicht erreicht werden kann und soll.

1.2 Grafische Darstellung

Die Box-Whisker-Diagramme sind für die vergleichende Darstellung von Merkmalen gut geeignet.

Die Box umfasst 50% der Merkmalswerte, die Whiskers stellen Maximum und Minimum dar. Statistiksoftware liefert meist noch den Median als Querstrich durch die Box und als Kreuzchen den arithmetischen Mittelwert.

Zusätzlich werden zwei Ausreißerwerte eingezeichnet, wobei für diese Werte außerhalb 1,5 Interquartilsabstand, dann gelten diese als ausreichend signifikant, um sie als Ausreißer zu bezeichnen.

Boxplot

Starplots stellen alle Werte eines multivariaten Datensatzes in normierter Sternform dar.

Der Vergleich von mehreren Datensätzen lässt das Streuverhalten der Variablen anschaulich beurteilen. So können z.B. periodische Änderungen auffallen, die wiederum Anlass zu Untersuchungen abgeben.

[Starplots](#)

Scatterplots stellen den paarweisen Zusammenhang zweier Variablen als xy-Diagramm dar. Statistiksoftware ordnet diese Plots analog der Korrelationsmatrix an. Im vorliegenden Fall wurde nur eine Hälfte der symmetrischen Matrix dargestellt.

Unterhalb der Plots sind der Korrelationskoeffizient sowie die Wahrscheinlichkeit aus dem t -Test für den Korrelationskoeffizienten.

Die Diagonale der Matrix ist mit den Varianzen der Variablen besetzt.

[Scatterplots](#)

1.3 Berechnungen

Zur vollständigen Korrelationsmatrix ist noch die Determinante und der Grad der Multikollinearität berechnet.

[Korrelationsmatrix](#)

2. EXCEL

Der Datenumfang ist aus Gründen der Anschaulichkeit bei den Scatterplots auf 500 Datensätze begrenzt.

Die Formeln sind für 10 Variable vorbereitet. Falls weniger Variable verwendet werden, sind die nicht zutreffenden Zellen mit einem leeren Text "" zugeordnet.

Teilweise erscheinen Fehlermeldungen in nicht belegten Zellen, die keine weiteren Auswirkungen haben.

Nullwerte sind, wo sinnvoll, ausgeblendet.

Eingaben werden, wo sinnvoll, mit der Funktion "Gültigkeit" überwacht.
Als Meldung wird nur eine Warnung ausgegeben. "Gültigkeit" kann auch total wirken!

Die Lösungsergebnisse der praxisorientierten EXCEL-Aufgaben für Variable, Faktoren, Verteilungen und Matrizen.
Die Ergebnisse sind in Excel-Dateien aufgeführt.
EXCEL-TIP > Die Liste der verwendeten Namen wird über die Funktionstaste F3 ausgegeben

Namen

Auf den Berechnungsblättern sind Tabellenfelder zur besseren Orientierung farblich unterlegt:

www.tqu-verlag.de

freie Eingabefelder

Beschriftungen

Ergebnisse (keine Einträge vornehmen), in der Regel gesperrt!

Die Gitternetzlinien und Spalten/Zeilenüberschriften sind auf den Berechnungsblättern der Übersichtlichkeit wegen ausgeblendet.

Die Boxplots folgen der Grafik für Aktienkurse (Eröffnung, Max, Min, geschlossen)
mit der Entsprechung für die Statistik (25% Quartil, Min, Max, 75% Quartil).

Die Starplots nutzen das Radardiagramm von EXCEL.

Die Grafiken greifen auf den Datenauszug zu, der die erforderlichen Zeilen bestimmt.

Die Scatterplots sind klassische XY-Diagramme.

Alle Diagramme sind in der Formatierung auf den Anwendungszweck angepasst.

Datenauszug

EXCEL-TIP > Zur Anwendung der Matrixfunktionen muss der jeweilige Zellbereich komplett markiert werden. Der Abschluss des Formelassistenten darf jetzt nicht durch den "OK" Button, sondern muss durch die Tastenkombination "**strg**", "**shift**", "**enter**" erfolgen

Als sichtbare Information wird die Formel durch EXCEL in der Eingabezeile in geschwungene Klammern gesetzt {}.
Diese Klammern können nicht manuell eingegeben werden.

Matrixfunktionen können nur noch für den gesamten zutreffenden Bereich geändert werden.

Im übrigen können auch "einfache" Funktionen zur Matrixfunktion aufgewertet werden und erreichen eine deutlich erweiterte Performance
(bspw. für die Berechnung von Eigenwerten oder die Berechnung von Eigenvektoren)

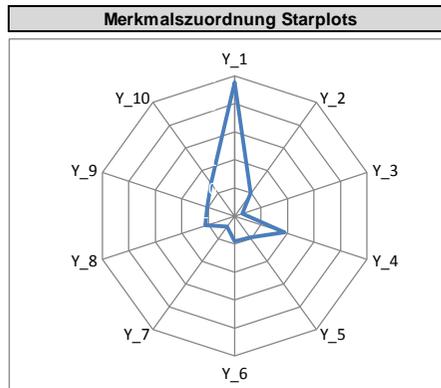
3. empfehlenswerte Literatur

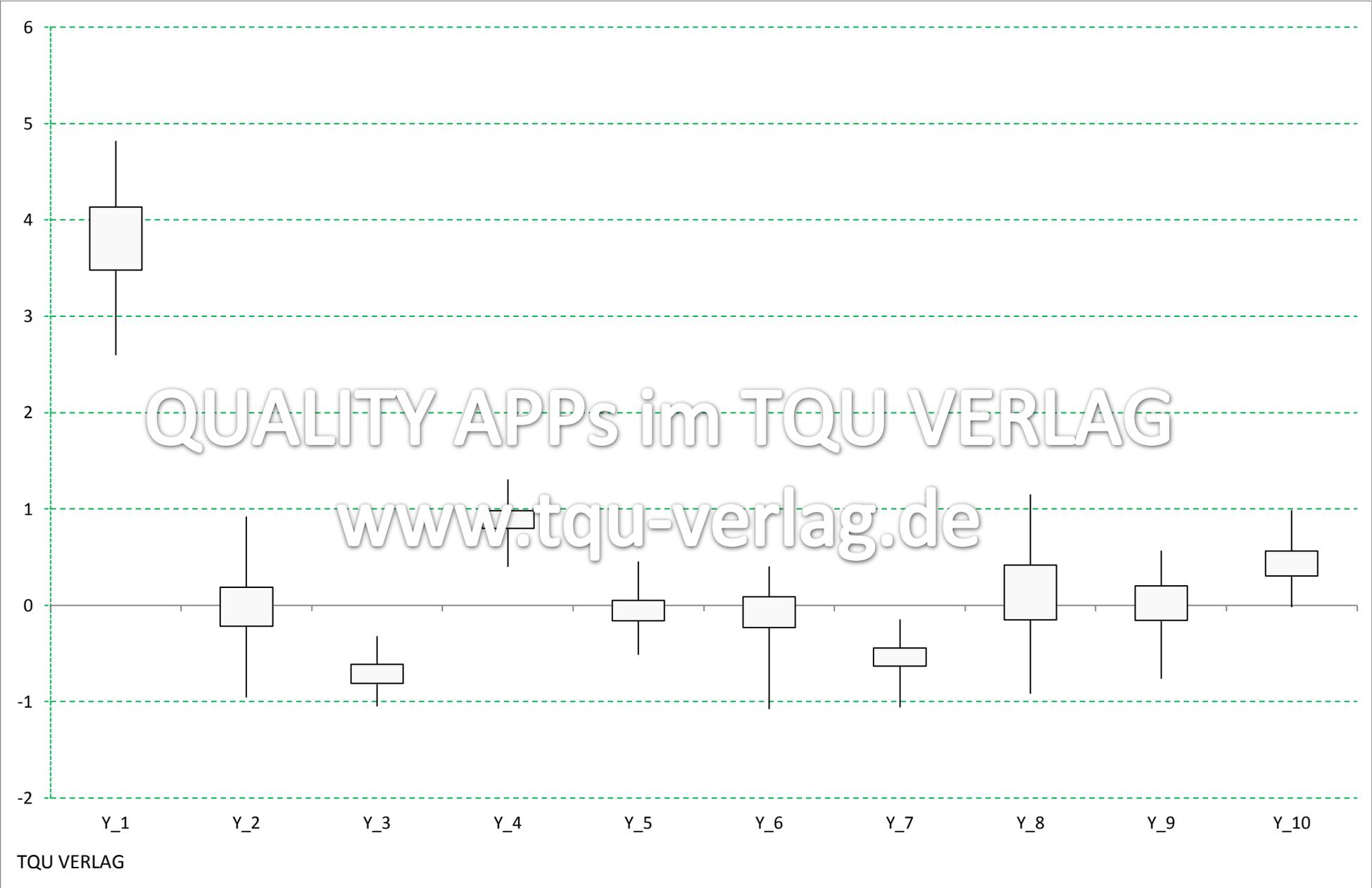
- [1] Jahn, W. Braun, L.: Praxis in der Qualität-Prozess-Management. Multivariate Statistik in 100 Beispielen. Carl Hanser Verlag, München 2006
- [2] Jahn, W. Reuter, K.: Komplex, aber beherrschbar. QZ (2011)5, S.34
- [3] Jahn, W. Reuter, K.: Workbook Einführung Multivariate Statistik TQU Verlag 2012
- [4] QUALITY APP "Multivariate Prozessverbesserung", TQU Verlag 2011
- [5] QUALITY APP "Multivariate_Prozessfähigkeit", TQU Verlag 2012
- [6] QUALITY APP "Multivariate Tolerierung", TQU Verlag 2012

Kunde	Automobil OEM
Artikel	Modell_A
Nummer	20
Bezeichnung	Säule_B

Variable	Benennung	Nummer	Einheit	Spezifikation			Statistik					
				LSL	USL	T	Mean	Stdev	25% Quantil	max	min	75% Quantil
Y_1	Maß1	20-01	mm				3,7687	0,4661	3,4775	4,8180	2,5940	4,1353
Y_2	Maß2	20-02					-0,0032	0,3045	-0,2170	0,9200	-0,9580	0,1853
Y_3	Maß3	20-03					-0,7177	0,1426	-0,8120	-0,3190	-1,0490	-0,6128
Y_4	Maß4	20-04					0,7881	0,1442	0,7988	1,3040	0,3970	0,9830
Y_5	Maß5	20-05					-0,0443	0,1625	-0,1593	0,4560	-0,5150	0,0518
Y_6	Maß6	20-06					-0,0866	0,2492	-0,2335	0,4030	-1,0780	0,0865
Y_7	Maß7	20-07					-0,5377	0,1575	-0,6333	-0,1490	-1,0620	-0,4420
Y_8	Maß8	20-08					0,1085	0,4152	-0,1530	1,1520	-0,9170	0,4180
Y_9	Maß9	20-09					0,0261	0,2472	-0,1583	0,5660	-0,7620	0,1998
Y_10	Maß10	20-10					0,4429	0,1915	0,3045	0,9880	-0,0190	0,5625

Datensätze	n	228
Freiheitsgrade	df	227
Variable	m	10





QUALITY APPs im TQU VERLAG

www.tqu-verlag.de

Auswahl Start						
1	1	2	3	4	5	6
Stichproben- umfang n						
	8	9	10			
						
	13	14	15	16	17	18
						

QUALITY APPs im TQU VERLAG
www.tqu-verlag.de

var	Y_1	Y_2	Y_3	Y_4	Y_5	Y_6	Y_7	Y_8	Y_9	Y_10
Y_1	3,79700									
		$r = -0,359$ $p = 0,00\%$	$r = -0,679$ $p = 0,00\%$	$r = 0,090$ $p = 17,44\%$	$r = 0,530$ $p = 0,00\%$	$r = 0,168$ $p = 1,09\%$	$r = -0,159$ $p = 1,59\%$	$r = 0,985$ $p = 0,00\%$	$r = 0,558$ $p = 0,00\%$	$r = -0,252$ $p = 0,01\%$
Y_2		0,08500								
			$r = 0,376$ $p = 0,00\%$	$r = 0,295$ $p = 0,00\%$	$r = -0,006$ $p = 93,35\%$	$r = -0,201$ $p = 0,22\%$	$r = 0,404$ $p = 0,00\%$	$r = -0,360$ $p = 0,00\%$	$r = -0,537$ $p = 0,00\%$	$r = -0,080$ $p = 22,87\%$
Y_3			-0,88900							
				$r = -0,036$ $p = 58,87\%$	$r = -0,237$ $p = 0,03\%$	$r = -0,171$ $p = 0,94\%$	$r = 0,252$ $p = 0,01\%$	$r = -0,677$ $p = 0,00\%$	$r = -0,416$ $p = 0,00\%$	$r = 0,506$ $p = 0,00\%$
Y_4				0,72100						
					$r = 0,133$ $p = 4,42\%$	$r = 0,275$ $p = 0,00\%$	$r = 0,112$ $p = 9,12\%$	$r = 0,088$ $p = 18,28\%$	$r = -0,030$ $p = 65,70\%$	$r = 0,019$ $p = 76,93\%$
Y_5					0,03700					
						$r = 0,043$ $p = 52,18\%$	$r = -0,015$ $p = 82,60\%$	$r = 0,398$ $p = 0,00\%$	$r = 0,234$ $p = 0,04\%$	$r = -0,053$ $p = 42,90\%$
Y_6						-0,10500				
							$r = -0,352$ $p = 0,00\%$	$r = 0,162$ $p = 1,40\%$	$r = 0,473$ $p = 0,00\%$	$r = 0,301$ $p = 0,00\%$
Y_7							-0,51900			
								$r = -0,181$ $p = 0,60\%$	$r = -0,355$ $p = 0,00\%$	$r = -0,162$ $p = 1,41\%$
Y_8								0,44100		
									$r = 0,548$ $p = 0,00\%$	$r = -0,262$ $p = 0,01\%$
Y_9									0,18700	
										$r = 0,373$ $p = 0,01\%$
Y_10										0,47400

QUALITY APPS im TQU-VERLAG
www.tqu-verlag.de

Korrelationsmatrix

		3	4	5	6	7	8	9	10		
	R	Y_1	Y_2	Y_3	Y_4	Y_5	Y_6	Y_7	Y_8	Y_9	Y_10
	Y_1	1	-0,359	-0,679	0,090	0,530	0,168	-0,159	0,985	0,558	-0,252
	Y_2	-0,3595	1	0,376	0,295	-0,006	-0,201	0,404	-0,360	-0,537	-0,080
3	Y_3	-0,679	0,376	1	-0,036	-0,237	-0,171	0,252	-0,677	-0,416	0,506
4	Y_4	0,090	0,295	-0,036	1	0,133	0,275	0,112	0,088	-0,030	0,019
5	Y_5	0,530	-0,006	-0,237	0,133	1	0,043	-0,015	0,398	0,234	-0,053
6	Y_6	0,168	-0,201	-0,171	0,275	0,043	1	-0,352	0,162	0,473	0,301
7	Y_7	-0,159	0,404	0,252	0,112	-0,015	-0,352	1	-0,181	-0,355	-0,162
8	Y_8	0,985	-0,360	-0,677	0,088	0,398	0,162	-0,181	1	0,548	-0,262
9	Y_9	0,558	-0,537	-0,416	-0,030	0,234	0,473	-0,355	0,548	1	0,373
10	Y_10	-0,252	-0,080	0,506	0,019	-0,053	0,301	-0,162	-0,262	0,373	1

Determinante R	0,0001342
Grad der Multikollinearität	7451

Signifikanz

p	Y_1	Y_2	Y_3	Y_4	Y_5	Y_6	Y_7	Y_8	Y_9	Y_10
Y_1		0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
Y_2			0,0000	0,0000	0,9335	0,0022	0,0000	0,0000	0,0000	0,2287
Y_3				0,5887	0,0003	0,0094	0,0001	0,0000	0,0000	0,0000
Y_4					0,0442	0,0000	0,0912	0,1828	0,6570	0,7693
Y_5						0,5218	0,8260	0,0000	0,0004	0,4290
Y_6							0,0000	0,0140	0,0000	0,0000
Y_7								0,0060	0,0000	0,0141
Y_8									0,0000	0,0001
Y_9										0,0000
Y_10										

Nr	Y_1	Y_2	Y_3	Y_4	Y_5	Y_6	Y_7	Y_8	Y_9	Y_10
1	3,797	0,532	-0,812	1,108	0,119	0,15	-0,554	0,122	-0,086	0,444
2	3,943	0,205	-0,64	0,87	-0,038	-0,078	-0,691	0,297	0,155	0,724
3	3,8	-0,723	-0,709	0,632	0,044	0,145	-0,659	0,016	0,389	0,796
4	3,355	0,085	-0,685	0,488	-0,198	0,375	-0,671	-0,182	-0,058	0,441
5	3,797	-0,958	-0,923	0,742	-0,08	0,256	-0,665	0,124	0,492	0,674
6	3,447	0,236	-0,952	0,795	-0,086	-0,472	-0,63	-0,117	-0,389	0,075
7	3,186	0,15	-0,889	0,786	-0,21	-0,352	-0,468	-0,382	-0,434	0,197
8	2,783	0,183	-0,636	0,574	-0,315	-0,154	-0,589	-0,651	-0,309	0,379
9	4,391	-0,056	-0,797	0,965	0,087	0,299	-0,668	0,689	-0,024	0,289
10	3,864	-0,045	-0,775	0,726	0,015	0,315	-0,628	0,136	0,26	0,515
11	4,316	-0,171	-0,79	0,887	0,094	-0,194	-0,419	0,483	0,326	0,634
12	4,539	-0,361	-0,989	0,9	0,027	0,136	-0,791	0,782	0,498	0,523
13	4,122	0,163	-0,897	1,088	0,037	0,403	-0,728	0,417	0,066	0,422
14	4,493	0,619	-0,928	0,9	0,014	0,253	0,375	0,761	0,12	0,359
15	4,6	-0,12	1,045	1,06	0,18	0,057	-0,54	0,475	0,2	0,397
16	4,346	-0,334	-0,974	0,817	0,003	-0,105	-0,808	0,63	0,279	0,389
17	4,189	0,122	-0,752	0,88	0,036	-0,088	-0,583	0,507	0,119	0,512
18	3,078	-0,138	-0,51	0,74	0,143	0,214	0,615	0,175	0,106	0,69
19	3,783	-0,21	-0,37	0,81	0,173	0,27	0,29	0,194	0,069	0,541
20	4,355	0,166	-0,848	0,91	-0,042	0,115	-0,601	0,733	0,295	0,551
21	4,492	-0,656	-0,919	0,997	0,264	0,171	-0,677	0,633	0,462	0,634
22	4,136	-0,259	-0,845	0,814	-0,058	-0,089	-0,537	0,441	0,166	0,307
23	4,748	0,204	-0,832	0,967	0,186	0,145	-0,686	0,993	-0,039	0,213
24	3,369	-0,088	-0,739	0,832	-0,057	0,122	-0,805	-0,323	-0,043	0,347
25	4,029	0,014	-0,701	0,761	0,109	0,191	-0,504	0,277	0,187	0,57
26	4,22	-0,361	-0,99	0,783	0,054	-0,226	-0,669	0,55	-0,288	-0,019
27	3,076	0,128	-0,58	0,914	-0,14	0,055	-0,89	-0,478	-0,158	0,537
28	3,945	-0,236	-0,768	0,725	0,051	0,198	-0,721	0,235	0,329	0,474
29	4,435	-0,145	-0,85	0,957	0,049	0,213	-0,516	0,66	0,563	0,75
30	4,212	-0,107	-0,82	1,022	-0,036	0,373	-0,47	0,555	-0,037	0,231
31	3,859	-0,105	-0,825	0,956	-0,125	0,206	-0,607	0,172	0,272	0,396
32	3,505	-0,336	-0,813	0,938	-0,131	-0,066	-0,558	-0,17	0,061	0,4
33	4,362	-0,137	-0,931	1,136	0,197	0,208	-0,541	0,542	0,183	0,457
34	4,423	-0,103	-0,854	0,924	0,073	0,279	-0,718	0,655	0,254	0,268