

TURKISH STATISTICAL INSTITUTE

TRAINING WORKSHOP ON RETAIL TRADE TURNOVER STATISTICS

**Cengiz ÖZKAN
22-24 MAY 2017
Muscat, Oman**

OUTLINE

➔ ***DATA COLLECTION***

➔ **SOURCE OF DATA**

➔ **MODE OF COLLECTION**

➔ **SAMPLING**

➔ **EDITING VALIDATION AND IMPUTATION**

DATA COLLECTION - SOURCE OF DATA

1- Monthly Retail Trade Questionnaire

2- Consumer Price Index (At COICOP level)

MONTHLY RETAIL TRADE QUESTIONNAIRE

- ➔ **Online questionnaire since January 2014**
- ➔ **Passwords distributed by Regional Offices**
- ➔ **Data entry until the end of following month**
- ➔ **Consistency checks during data entry**
- ➔ **Data analyses in Regional Offices and Central Office**

CONSUMER PRICE INDEX (AT COICOP LEVEL)

- ➔ Deflator for volume index
- ➔ The CPI at COICOP (Classification of Individual Consumption According to Purpose) level is converted to NACE Rev.2 level
- ➔ Household Budget Survey data is used for this conversion

COICOP	2014												2015		
	OCAK	SUBAT	MART	NISAN	MAYIS	HAZİRAN	TEMMUZ	AĞUSTOS	EYLÜL	EKİM	KASIM	ARALIK	OCAK	SUBAT	MART
4 0111 EKMEK VE TAHILLAR	253,62	259,15	264,72	268,63	272,48	275,62	277,73	279,10	280,01	282,71	286,34	287,81	289,19	290,34	290,48
4 0112 ET	255,69	263,15	266,41	268,73	265,25	268,55	276,71	280,34	283,16	280,36	279,19	279,39	285,37	294,33	300,10
4 0113 BALIK	217,86	226,75	232,90	227,56	254,83	245,70	244,37	257,58	228,76	223,64	220,58	236,18	255,84	266,03	268,42
4 0114 SÜT, PEYNİR VE YUMURTA	238,10	241,52	243,47	241,40	238,91	244,03	245,70	255,21	261,37	261,39	260,12	263,14	265,58	265,82	266,65
4 0115 KATI VE SIVI YAĞLAR	214,23	220,22	223,52	225,76	226,57	226,78	227,38	228,23	230,05	231,30	232,13	233,02	239,92	247,13	250,85
4 0116 MEYVELER	204,91	213,32	226,18	248,68	280,44	285,89	266,90	254,04	244,31	245,29	238,23	231,00	240,44	248,67	267,67
4 0117 SEBZE	347,71	317,97	326,68	324,27	268,79	254,90	259,53	267,05	272,48	316,55	322,22	320,10	363,43	389,78	413,14
4 0118 ŞEKER, REÇEL, BAL,	200,32	202,80	204,07	206,89	208,34	212,53	215,27	216,48	217,18	218,66	220,19	220,70	221,50	222,48	224,98
4 0119 BAŞKA YERDE SINIFLA	258,64	260,75	263,61	267,11	268,57	271,03	269,99	272,93	273,60	276,21	277,77	278,73	280,47	283,43	285,23
4 0121 KAHVE, ÇAY VE KAKA	235,09	235,95	238,21	241,15	242,08	247,65	253,85	258,40	261,39	265,92	270,81	273,95	276,24	281,40	284,85
4 0122 MADEN SUYU, ALKOLS	188,69	190,45	191,35	191,71	192,88	194,07	194,77	196,32	197,15	198,57	198,69	197,94	198,03	199,16	199,83
4 0211 İÇKİLER	450,02	450,49	449,46	450,03	449,99	448,98	459,61	466,28	466,72	466,64	466,23	465,08	506,45	507,44	506,90
4 0212 ŞARAP	286,64	283,33	273,42	277,34	275,90	266,60	274,48	288,37	277,83	277,49	278,26	275,57	286,21	288,65	287,54
4 0213 BİRA	389,96	392,98	393,03	394,90	392,57	394,31	401,82	411,92	413,19	413,05	413,63	413,33	441,54	443,75	443,59
4 0220 TÜTÜN	411,95	411,95	411,95	411,95	411,39	411,16	419,31	416,63	411,65	411,65	411,65	411,65	427,89	425,72	425,72
4 0311 KUMAŞLAR	191,47	195,24	198,00	199,24	200,23	200,88	201,15	201,54	201,93	202,45	203,57	204,41	205,03	206,36	207,57
4 0312 GİYSİLER	141,72	134,22	136,46	156,39	174,39	171,45	165,25	156,35	150,06	166,64	174,44	166,60	154,16	143,63	144,03
4 0313 DİĞER GİYSİ ÇEŞİTL	159,95	156,98	160,70	173,40	177,58	177,48	175,62	171,53	173,50	182,43	185,59	183,03	176,60	171,25	175,24
4 0314 KURU TEMİZLEME, Gİ	232,65	234,79	237,50	239,17	240,35	242,32	244,39	246,33	248,33	251,29	253,02	254,00	256,40	257,42	258,93
4 0321 AYAKKABI VE DİĞER	163,74	154,59	162,43	180,48	188,23	187,73	182,17	174,03	176,62	191,76	197,72	193,52	180,30	167,45	172,92
4 0322 AYAKKABI ONARIMI V	263,35	264,98	267,73	270,89	273,64	276,71	278,27	278,92	279,75	281,84	285,39	289,19	291,67	294,26	296,86
4 0410 KİRACI TARAFINDAN	317,12	318,48	319,67	321,13	323,04	325,48	327,65	329,92	332,79	335,33	336,95	338,72	340,63	342,26	343,69
4 0431 KONUTUN BAKIM VE O	190,59	194,51	197,00	198,18	199,89	200,86	201,75	202,54	203,36	204,35	205,25	205,99	207,38	208,96	210,83
4 0441 SU (Şebeke Suyu)	227,99	230,56	231,35	232,89	237,35	238,16	242,14	243,88	244,25	247,86	250,90	250,31	267,79	270,84	272,29
4 0451 ELEKTRİK	233,01	233,01	233,01	233,73	233,73	233,73	233,79	233,79	233,79	254,90	254,90	254,90	254,51	254,51	254,51
4 0452 GAZ	327,48	325,20	324,02	322,34	319,09	318,46	320,24	319,66	318,99	333,30	327,71	325,16	320,32	319,54	326,86
4 0454 KATI YAKITLAR	270,20	272,37	271,95	270,55	269,30	266,49	265,13	264,05	269,73	272,42	275,38	277,79	282,14	285,21	287,55
4 0511 MOBİLYA VE MEFRUŞA	218,43	220,76	220,63	228,14	227,76	229,44	228,01	226,53	226,90	227,74	229,66	230,43	236,90	236,55	238,60
4 0512 HALI VE DİĞER YER	194,85	197,68	199,65	203,76	205,34	206,51	204,16	205,14	204,79	207,69	209,08	209,39	210,41	211,37	212,18
4 0520 EV İÇİ TEKSTİL	177,01	179,62	181,20	186,04	186,74	187,64	188,02	188,26	188,79	185,75	185,74	185,26	188,47	188,76	190,61
4 0531 ELEKTRİKLİ VE ELEK	126,05	128,27	130,26	130,61	128,79	128,94	130,47	131,79	132,06	131,87	133,55	133,23	137,12	137,55	137,00
4 0532 ELEKTRİKLİ KÜÇÜK E	146,61	148,55	150,00	150,07	146,79	146,88	148,21	148,69	150,16	151,69	153,63	154,08	161,07	161,34	162,83
4 0533 EV ALETLERİNİN ONA	294,34	298,25	303,38	302,93	299,52	298,09	298,36	301,16	311,71	315,67	317,05	317,83	332,29	337,00	337,71
4 0540 ZÜCCACİYE, YEMEK T	191,74	194,08	196,37	198,45	200,56	202,10	203,77	205,42	206,88	208,94	210,35	211,69	212,23	213,83	216,13
4 0552 KÜÇÜK ARAÇLAR VE Ç	213,74	217,01	218,80	219,97	221,33	221,93	223,73	224,18	225,90	227,52	229,04	230,52	230,93	232,02	233,49
4 0561 DAYANIKSIZ EV MALZ	147,48	148,79	151,18	152,41	153,16	153,37	153,36	153,94	154,74	155,65	156,36	157,20	157,72	158,26	159,54
4 0562 EV İŞİ VE HANEHALK	346,05	352,00	355,22	356,23	357,54	360,94	364,19	368,68	371,20	371,26	377,47	374,84	381,38	382,95	386,22
4 0611 İLAÇLAR	93,46	93,59	93,35	93,79	95,59	98,51	98,88	98,88	98,88	98,88	98,88	98,88	98,88	100,13	101,22
4 0612 DİĞER SAĞLIK ÜRÜN	157,31	157,71	159,46	160,99	161,43	162,33	163,19	163,81	164,76	166,59	168,26	168,85	172,25	172,83	173,90

MODE OF COLLECTION

- ➔ **Web Based Computer Software**
- ➔ **Harzemli**
- ➔ **Data Collection And Processing**
- ➔ **DDI + Rule Files + Classification Server**

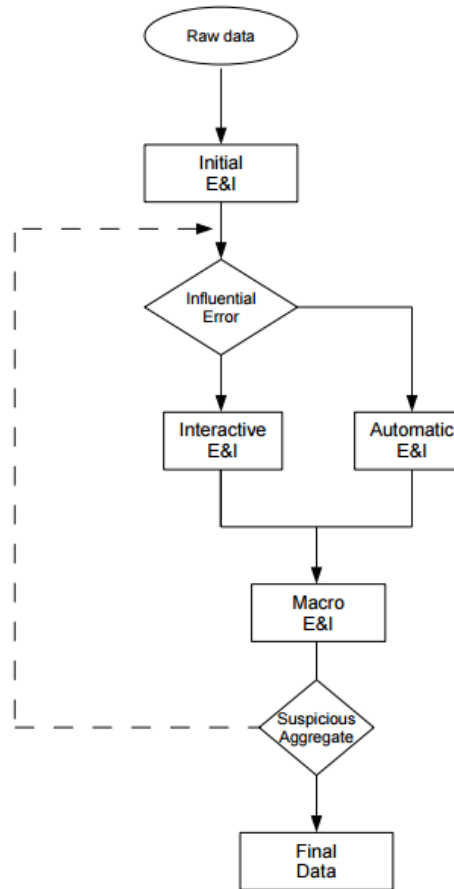
SAMPLING

- ➔ **Weights - Annual Industry and Services Statistics.**
- ➔ **Stratified random sampling**
- ➔ **Aggregation:**
 - ➔ **Food, drinks and tobacco**
 - ➔ **Non-food (except automotive fuel)**
 - ➔ **Automotive fuel**
 - ➔ **Retail Trade in general.**

SAMPLING

SIZE GROUP	NUMBER OF ENTERPRISE	SAMPLING STATUS
0-4	276	STRATIFIED RANDOM SAMPLING
5-49	662	STRATIFIED RANDOM SAMPLING
50 +	792	CENCUS
TOTAL	1727	

DATA VALIDATION & EDITING



DATA VALIDATION & EDITING

- ➔ **E&I strategy:**
 - ➔ **1. identify and eliminate**
 - ➔ **2. influential errors**
 - ➔ **3. final output**

DATA VALIDATION & EDITING

➔ **Four principles of the E&I process**

- 1. Data quality**
- 2. Control of the process.**
- 3. Needs of the users.**
- 4. Simple, cheap and fast**

TURKSTAT PRACTICE-DATA VALIDATION-EDITING

- ➔ **RULES – WEB BASED QUESTIONNAIRE ENTRY**
- ➔ **STANDARD PERSPECTIVE – ENTERPRISES**
- ➔ **ANALYSE PERSPECTIVE – REGIONAL OFFICES**

ANALYSES OF THE REGIONAL OFFICES

- ➔ **Automated data analysis procedures**
 - ➔ **SAS reports**
 - ➔ **Possible erroneous data is queried**
 - ➔ **Correction or validation of data**

ANALYSES OF THE CENTRAL OFFICES

- ➔ **Automated and manual data analysis procedures**
 - ➔ **SAS reports**
 - ➔ **Deeper analyses for more important data**
 - ➔ **Analyses at index level**
 - ➔ **Possible erroneous data is queried and reported to Regional Offices**
 - ➔ **Correction or validation of data**

IMPUTATION

- Apply growth rate for actual month for units in same industry and size group

$$T_t = T_{t-1} \cdot \sum_{i \in M} T_t^i / \sum_{i \in M} T_{t-1}^i,$$

Where M is the matched sample set of units from same industry and size group, responding in both t and $t - 1$.

IMPUTATION

ADVANTAGES

- ➔ Easy to apply
- ➔ Unaffected by structural changes in previous year
- ➔ Take into account trend changes among similar enterprises
- ➔ Can be used if there is no data on the unit t-12 or t-13
- ➔ Can be used in religious festivals

IMPUTATION

DISADVANTAGES

- ➔ Does not take into account units' special seasonal pattern
- ➔ Affected by outliers/errors among similar units
- ➔ A sort of “mean imputation”, which can be bad!

OUTLIER IDENTIFICATION

- ➔ **Box-Plot – Micro level**
- ➔ **T-ERROR – Index level**

BOX-PLOT

- ➔ Lower and Upper Outer Fences (LOF and UOF, respectively) with $k=3$ are calculated:

$$LOF = Q_1 - k(Q_3 - Q_1)$$

$$UOF = Q_3 + k(Q_3 - Q_2)$$

with

$$k = 3$$

- ➔ Any observation outside of the outer fences is marked as problematic outlier.

BOX-PLOT

- ➔ **With the box-plot method the observations of a time series are rearranged into quartiles**
- ➔ **Outliers are detected for each strata: Activity and size class**
- ➔ **Based on the 1st and 3rd quartile lower and upper fences are determined**
- ➔ **Observations beyond the fences are potential outliers or statistical breaks**

BOX-PLOT

- ➔ **Why box-plot for determining outliers?**
 - ➔ **Easy to use and understand**
 - ➔ **It is quite transparent**
 - ➔ **Does not demand any assumptions about the distribution of the data or the number of observations**
 - ➔ **The outliers do not take part in the calculation of the control limits**
 - ➔ **Easily handles large data sets**

T-Error

- ➔ **TERROR (TRAMO for errors) is an application to quality control of data**
- ➔ **In particular, to the detection of errors in reported indices data**
- ➔ **For each series, the program automatically identifies an ARIMA model and detects several types of outliers**

Seats / Tramo

Series Model Models Run Output Out Tables Out Matrix Graph Load Save Fame About

Series Name

Series Attributes

Observations Starting Year

First Period #Obs / Period

Iter Parameter

Iter = 0 Iter = 1

Iter = 2 Iter = 3

Seats/Tramo

Seats Tramo/Seats

Tramo Error

T-Error

Terror TSW Series List

DATE : 2017-05-04 17:16:58

Input Parameters :

mq=12 out= 0 int2= 0 rsa= 5 terror= 1 sens= 0 units=-1 modelsum= 1 k1= 3.000 k2= 4.000

Summary Statistics

18 Series were tested.

0 Releases were suspicious (possibly wrong).

0 Releases were very suspicious (likely wrong).

0 Series produced a Run-Time EXCEPTION.

0 Series did not match TERROR memory constraints.

18 Series passed the plausibility tests.

THANK YOU