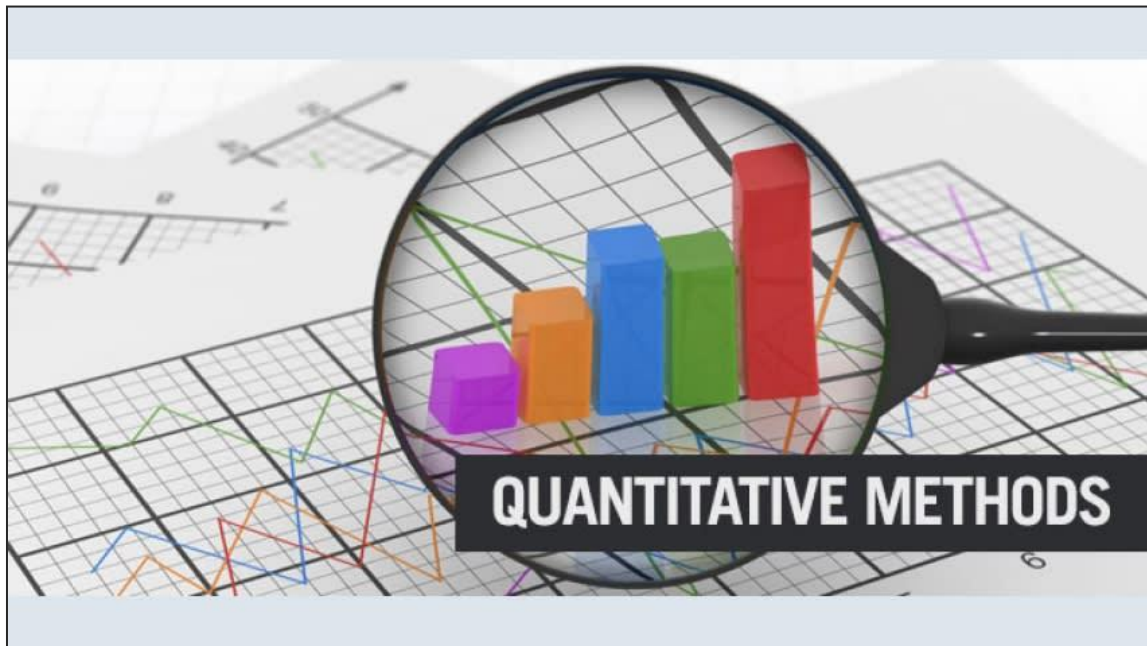


Ποσοτικές Μέθοδοι για Οικονομικές Αποφάσεις

Ομαδική Εργασία 3

ΜΠΟΥΦΙ ΓΙΩΑΝ ΜΟΕΣ/1931
ΚΑΣΦΙΚΗΣ ΚΩΝΣΤΑΝΤΙΝΟΣ
ΑΝΤΩΝΑΙΟΣ ΔΗΜΗΤΡΗΣ
ΜΑΝΕΤΤΑΣ ΙΩΑΝΝΗΣ ΜΟΕΣ/1926
12/30/2019



[Type the abstract of the document here. The abstract is typically a short summary of the contents of the document. Type the abstract of the document here. The abstract is typically a short summary of the contents of the document.]

Περιεχόμενα

Ερώτημα 1	3
Ερώτημα 2	7
Ερώτημα 3	8
Ερώτημα 4	10
Ερώτημα 5	12
Ερώτημα 6	13
Ερώτημα 7	15
Bonus Ερώτημα	18

Ερώτημα 1 - Descriptives

Descriptive Statistics

	N	Range	Minimum	Maximum	Sum	Mean
	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic
Employee's Age in 1980	50	47	22	69	1846	36,92
Yearly Salary in 1982	40	41048	9652	50700	759389	18984,73
Increase in Salary over 1981	39	8125	0	8125	96483	2473,92
Valid N (listwise)	39					

One-Sample Test

	Test Value = 0					
	t	df	Sig. (2-tailed)	Mean Difference	99% Confidence Interval of the Difference	
					Lower	Upper
Employee's Age in 1980	22,283	49	,000	36,920	32,48	41,36
Yearly Salary in 1982	12,229	39	,000	18984,725	14781,03	23188,42
Increase in Salary over 1981	8,936	38	,000	2473,923	1723,25	3224,60

Descriptive Statistics

	Mean	Std. Deviation	Variance	Skewness	
	Std. Error	Statistic	Statistic	Statistic	Std. Error
Employee's Age in 1980	1,657	11,716	137,259	1,061	,337
Yearly Salary in 1982	1552,376	9818,086	96394810,204	1,698	,374
Increase in Salary over 1981	276,842	1728,876	2989011,283	1,212	,378
Valid N (listwise)					

Descriptive Statistics

	Kurtosis	
	Statistic	Std. Error
Employee's Age in 1980	,397	,662
Yearly Salary in 1982	3,003	,733
Increase in Salary over 1981	1,869	,741
Valid N (listwise)		

Statistics				
		Increase in Salary over 1981	Yearly Salary in 1982	Employee's Age in 1980
N	Valid	39	40	50
	Missing	11	10	0
Median		2165,00	16360,50	34,50
Percentiles	25	1490,00	11857,00	27,75
	50	2165,00	16360,50	34,50
	75	3250,00	25378,75	42,75

Από την άνω περιγραφική στατιστική ανάλυση προκύπτει ότι :

1. Η **μέση τιμή** ή αλλιώς η κεντρική τάση των μεταβλητών :

- Μεταβλητή Age: 36.92
- Μεταβλητή salary82: 18984.73
- Μεταβλητή raise82: 2473.92

Τα υπολογισθέντα διαστήματα εμπιστοσύνης ($\alpha=1\%$), δηλαδή το διάστημα τιμών μέσα στο οποίο αναμένουμε να βρίσκεται η πληθυσμιακή μέση τιμή ανά μεταβλητή, είναι :

- Μεταβλητή age: $32.48 < X < 41,36$
- Μεταβλητή salary: $8214781.03 < X < 23188,42$
- Μεταβλητή raise82: $1723.25 < X < 3224.60$

2. Η **διάμεσος των μεταβλητών**, δηλαδή ο αριθμός που βρίσκεται ακριβώς στη μέση, έτσι ώστε το 50 % των ταξινομημένων αριθμών να είναι πάνω από τη διάμεσο είναι:

- Μεταβλητή Age: 34.50
- Μεταβλητή salary82: 16360.50
- Μεταβλητή raise82: 2165

Το πρώτο πράγμα που παρατηρούμε είναι ότι οι τιμές της διαμέσου είναι μικρότερες από αυτές της μέσης τιμής. Επίσης όσον αφορά την επιχείρηση γίνεται εμφανές ότι το προσωπικό της είναι νεαρής ηλικίας καθώς οι μισοί εργαζόμενοι είναι κάτω των 34.5 ετών.

3. Οι **μέγιστες** και **ελάχιστες τιμές** των δεδομένων που έχουν συγκεντρωθεί όσον αφορά τις τρεις μεταβλητές είναι :

- Μεταβλητή Age: 22(min) – 69(max)
- Μεταβλητή salary82: 9652(min) – 50700(max)

- Μεταβλητή raise82: $0(min) - 8125(max)$

Οι μισθοί και η αύξηση μισθού στην συγκεκριμένη επιχείρηση κατά πάσα πιθανότητα παρουσιάζουν ανισοκατανομή ή ακραίες τιμές.

4. Το **εύρος** αποτελεί μέτρο διασποράς το οποίο προκύπτει με την αφαίρεση της μικρότερης τιμής από την μεγαλύτερη ($max - min$). Το εύρος των μεταβλητών μας είναι :

- Μεταβλητή Age: 49
- Μεταβλητή salary82: 41084
- Μεταβλητή raise82: 8125

Γενικά το εύρος δεν είναι αποδεκτό, καθώς δεν αποτελεί αξιόπιστο μέτρο διασποράς.

5. Τα επικρατέστερα μέτρα διασποράς είναι η **τυπική απόκλιση** και η **διακύμανση**. Η τυπική απόκλιση στην αύξηση μισθού(1728) αν λάβουμε υπόψιν ότι η μέση τιμή της συγκεκριμένης μεταβλητής είναι 2473 είναι πολύ μεγάλη. Το παραπάνω δικαιολογείται καθώς τα δεδομένα έχουν ελάχιστη τιμή το μηδέν, δηλαδή ορισμένοι εργαζόμενοι δεν έχουν πάρει αύξηση. Η τυπική απόκλιση όσον αφορά την ηλικία σε συνδυασμό με την διάμεσο και μέση τιμή μας επιβεβαιώνει ότι η επιχείρηση έχει νέους σε ηλικία εργαζομένων. Στην περίπτωση ακραίων τιμών, όπως προαναφέρθηκε, η τυπική απόκλιση, η οποία βασίζεται στην μέση τιμή, δεν αποτελεί αξιόπιστο μέτρο διασποράς. Σε αυτήν την περίπτωση χρησιμοποιείται το ενδοτεταρτημοριακό εύρος.
6. Το **ενδοτεταρτημοριακό εύρος (IQR: interquartile range)** είναι η διαφορά μεταξύ του 3ου (Q3) και του 1ου (Q1) τεταρτημόριου,

$$IQR = Q3 - Q1$$

Για τις μεταβλητές μας οι τιμές που προκύπτουν είναι οι εξής :

- Μεταβλητή Age: $42.75 - 27.72 = 15.03$
- Μεταβλητή salary82: $25378.75 - 11857 = 13521.75$
- Μεταβλητή raise82: $3250 - 1490 = 1760$

7. **Λοξότητα-Κυρτότητα** : Και οι τρεις μεταβλήτες έχουν θετική λοξότητα(Skewness)το οποίο σημαίνει ότι υπάρχει ασυμετρία στην κατανομή των δεδομένων ως προς τα δεξιά. Η κύρτωση(Kurtosis) είναι επίσης θετική και στις τρεις μεταβλήτες, η μεταβλητή age είναι έχει την μικρότερη κυρτοσή(<1) η οποία όμως είναι υψηλότερη από της

κανονικής κατανομής που σημαίνει ότι έχει «λεπτές ουρές». Η μεταβλητή $raise_{82}$ παρουσιάζει μέτρια προς υψηλή κυρτόση με λεπτές ουρές και τέλος η μεταβλητή $salary_{82}$ παρουσιάζει ιδιαίτερα υψηλή κυρτόση με πολύ λεπτές ουρές. Καταληγούμε στο συμπέρασμα ότι τα δεδομένα μας δεν έχουν κανονική κατανομή καθώς τόσο η λοξότητα όσο και η κύρτωση έχουν υψηλές τιμές.

Ερώτημα 2 - Stem Leaf Plots

Μεταβλητή « Yearly Salary in 1982 »

Yearly Salary in 1982 Stem-and-Leaf Plot

Frequency	Stem &	Leaf
4,00	0 .	9999
13,00	1 .	0001112223344
11,00	1 .	56666777799
2,00	2 .	14
5,00	2 .	55566
2,00	3 .	12
1,00	3 .	5
2,00	Extremes	(>=48100)

Stem width: 10000
Each leaf: 1 case(s)

Μεταβλητή « Increase in Salary over 1981 »

Increase in Salary over 1981 Stem-and-Leaf Plot

Frequency	Stem &	Leaf
3,00	0 .	000
3,00	0 .	689
4,00	1 .	0234
9,00	1 .	555667789
6,00	2 .	122334
3,00	2 .	569
3,00	3 .	222
1,00	3 .	8
1,00	4 .	1
2,00	4 .	57
2,00	5 .	24
1,00	5 .	8
1,00	Extremes	(>=8125)

Stem width: 1000
Each leaf: 1 case(s)

Το Stem Leaf Plot αποτελεί διάγραμμα όπου κάθε τιμή που αναπαρίσταται χωρίζεται σε leaf που αναπαριστά συνήθως τα τελευταία ψηφία της εκάστοτε τιμής και σε stem όπου αναπαριστά τα πρώτα ψηφία της εκάστοτε τιμής. Στο πρώτο διάγραμμα το πλάτος τους stem είναι 10000 ενώ στο δεύτερο 1000. Για παράδειγμα η 4η γραμμή του πρώτου διαγράμματος αφορά 2 εργαζόμενους (frequency) όπου ο πρώτος έλαβε ετήσιες απολαβές:

$$10000 \text{ (stem width)} * 2 \text{ (stem)} + 1000 * 1 \text{ (leaf)} = 21000 \text{ ευρώ}$$

Ενώ ο δεύτερος έλαβε

$$10000 \text{ (stem width)} * 2 \text{ (stem)} + 1000 * 4 \text{ (leaf)} = 24000 \text{ ευρώ}$$

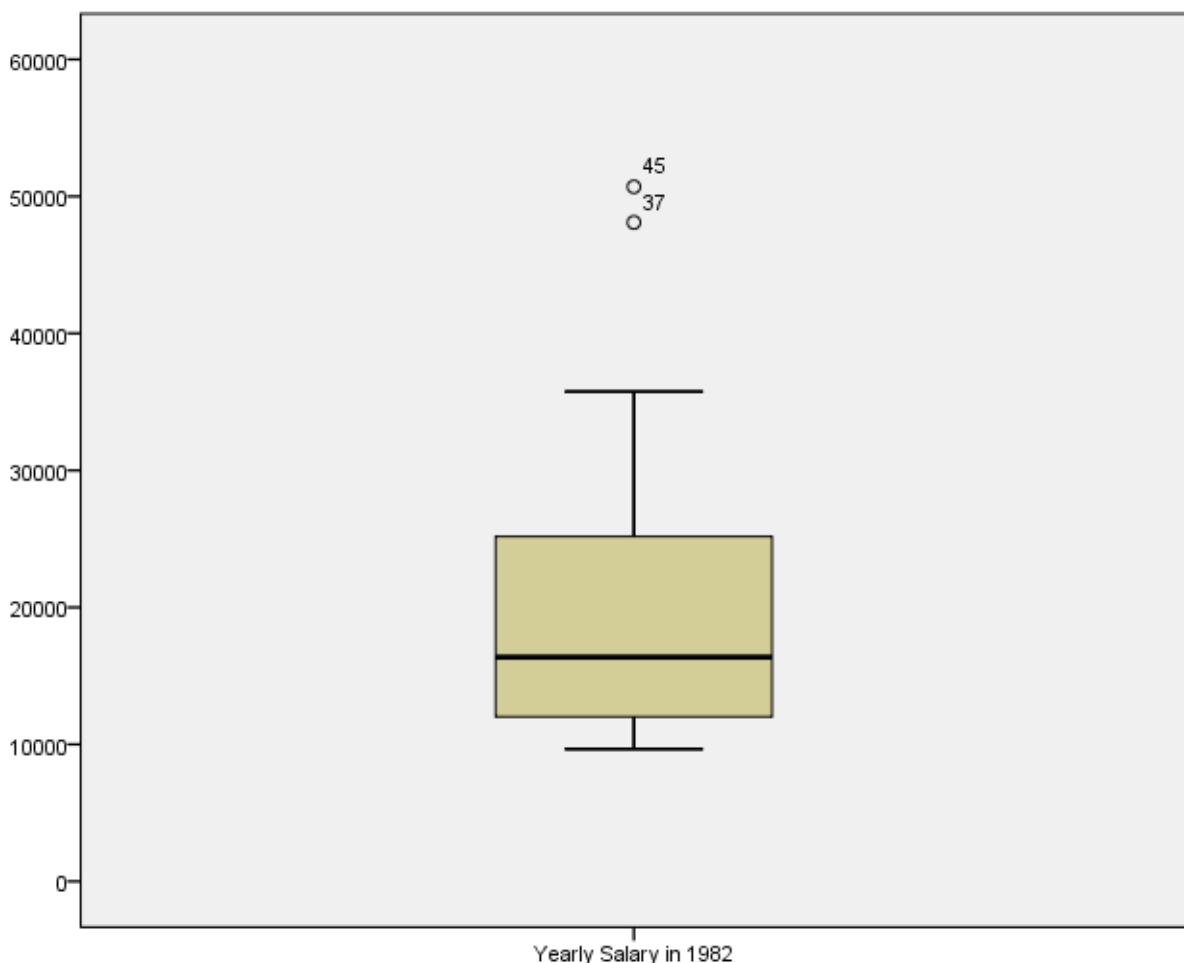
Στα παραπάνω διαγράμματα η ακρίβεια των τιμών είναι μεγαλύτερη - ίση του 1000.

Ερώτημα 3 - Box Plot

Το boxplot είναι ένας τυποποιημένος τρόπος απεικόνισης της κατανομής των δεδομένων βάσει των εξής :

- ελάχιστο
- πρώτο τεταρτημόριο (Q1)
- μέση τιμή
- τρίτο τεταρτημόριο (Q3)
- μέγιστο

Μεταβλητή « Yearly Salary in 1982 »



Από το παραπάνω διάγραμμα με την κάθετη γραμμή εντός του σκιασμένου πλαισίου μπορεί να εξαχθεί η μέση τιμή των ετήσιων απολαβών των εργαζομένων της εταιρίας (18984 ευρώ).

Με την κάθετη γραμμή κάτω από το σκιασμένο πλαίσιο μπορεί να εξαχθεί η ελάχιστη τιμή του δείγματος δεδομένων (9.652 ευρώ). Η κάθετη γραμμή του ελάχιστου πρέπει να είναι μεγαλύτερη-ίση του:

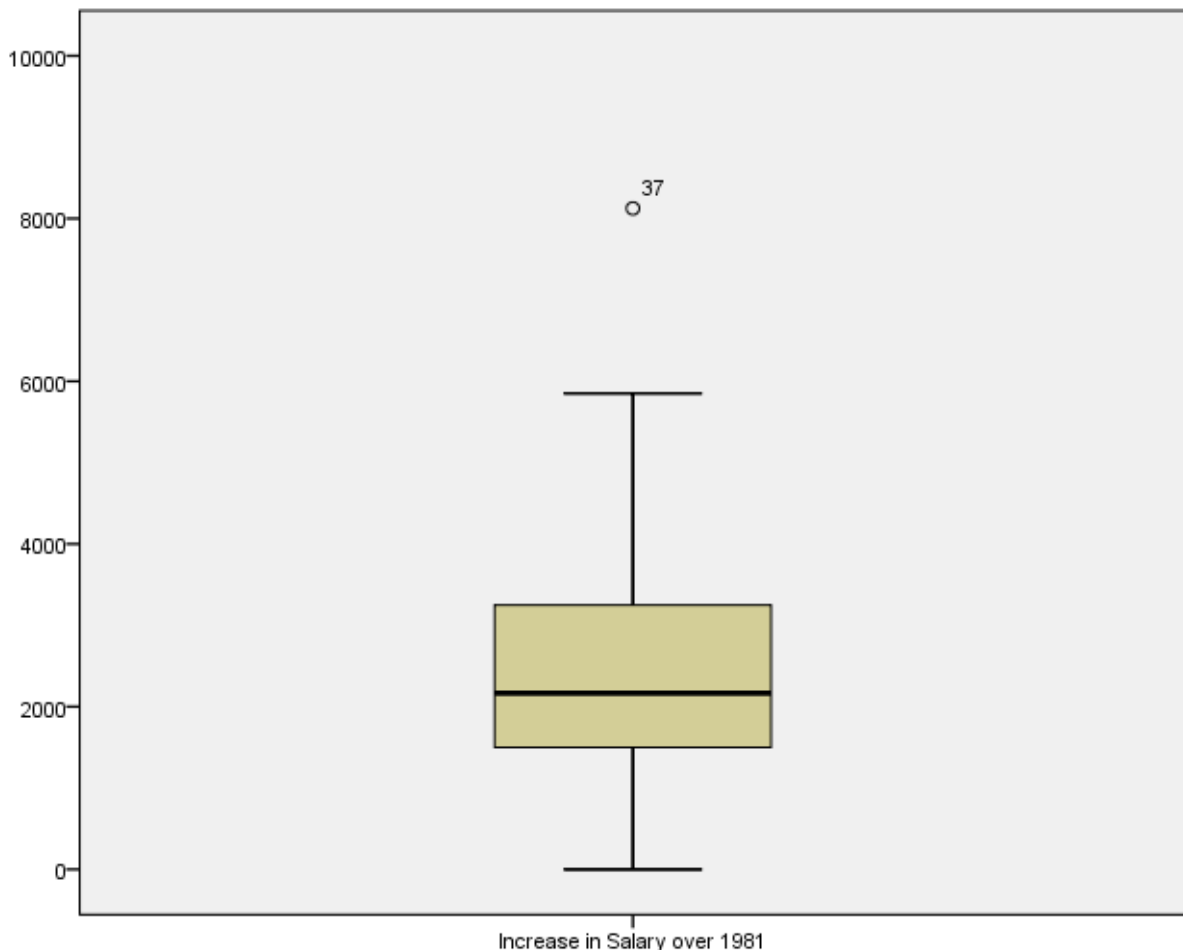
$$Q1 - (1.5 * IQR)$$

Αντίστοιχα βάσει της κάθετης γραμμής επάνω από το σκιασμένο πλαίσιο μπορεί να εξαχθεί η μέγιστη τιμή του δείγματος δεδομένων (35000 ευρώ). Η κάθετη γραμμή του μέγιστου πρέπει να είναι μικρότερη-ίση του:

$$Q3 + (1.5 * IQR)$$

Τιμές που είναι μικρότερες της ελάχιστης κάθετης γραμμής ή μεγαλύτερες της μέγιστης κάθετης γραμμής αναφέρονται ως outliers (extremes) και αποτυπώνονται με την χρήση σημείων (εγγραφές με σειρά αρίθμησης 37 και 45). Η σκιασμένη περιοχή αποτυπώνει το ενδοτεταρτημοριακό εύρος (IQR) που αποτελεί την διαφορά μεταξύ 3^{ου} και 1^{ου} τεταρτημόριου.

Μεταβλητή « Increase in Salary over 1981 »



Αντίστοιχα για το box plot που αναπαριστά την αύξηση στις ετήσιες απολαβές των εργαζομένων μπορούν να εξαχθούν τα εξής δεδομένα :

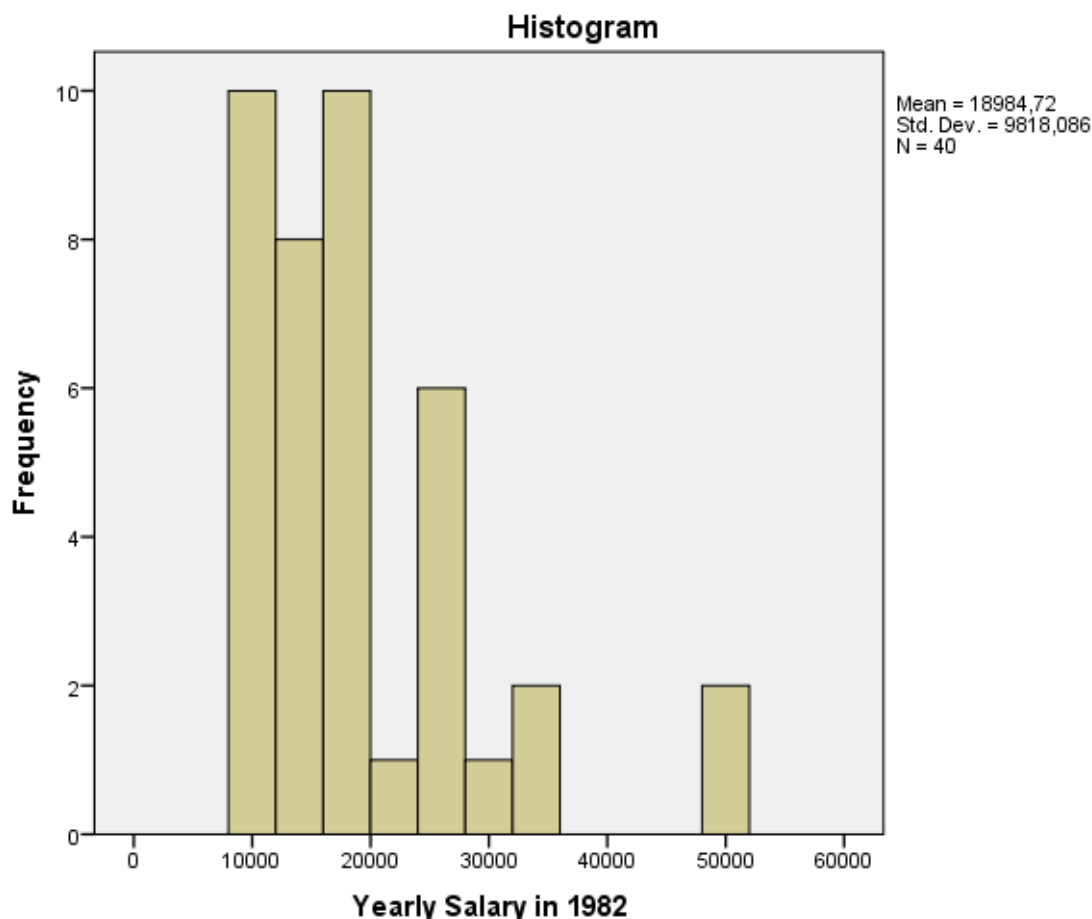
- Μέση τιμή : 2474 ευρώ
- Ελάχιστη τιμή : 0 ευρώ
- Μέγιστη τιμή : 5800 ευρώ
- Outlier : Εγγραφή με αρίθμηση 37

Ενώ με το σκιασμένο πλαίσιο αναπαρίσταται και εδώ το ενδοτεταρτημοριακό εύρος

Ερώτημα 4 - Ιστογράμματα

Το ιστόγραμμα σχηματίζεται από παρακείμενα ορθογώνια. Η επιφάνεια κάθε ορθογωνίου είναι μέτρο της συχνότητας εμφάνισης της συγκεκριμένης περιοχής τιμών ενώ το ύψος του ισούται με το λόγο της συχνότητας προς το εύρος των τιμών που αντιπροσωπεύει το ορθογώνιο.

Μεταβλητή «Yearly Salary in 1982»

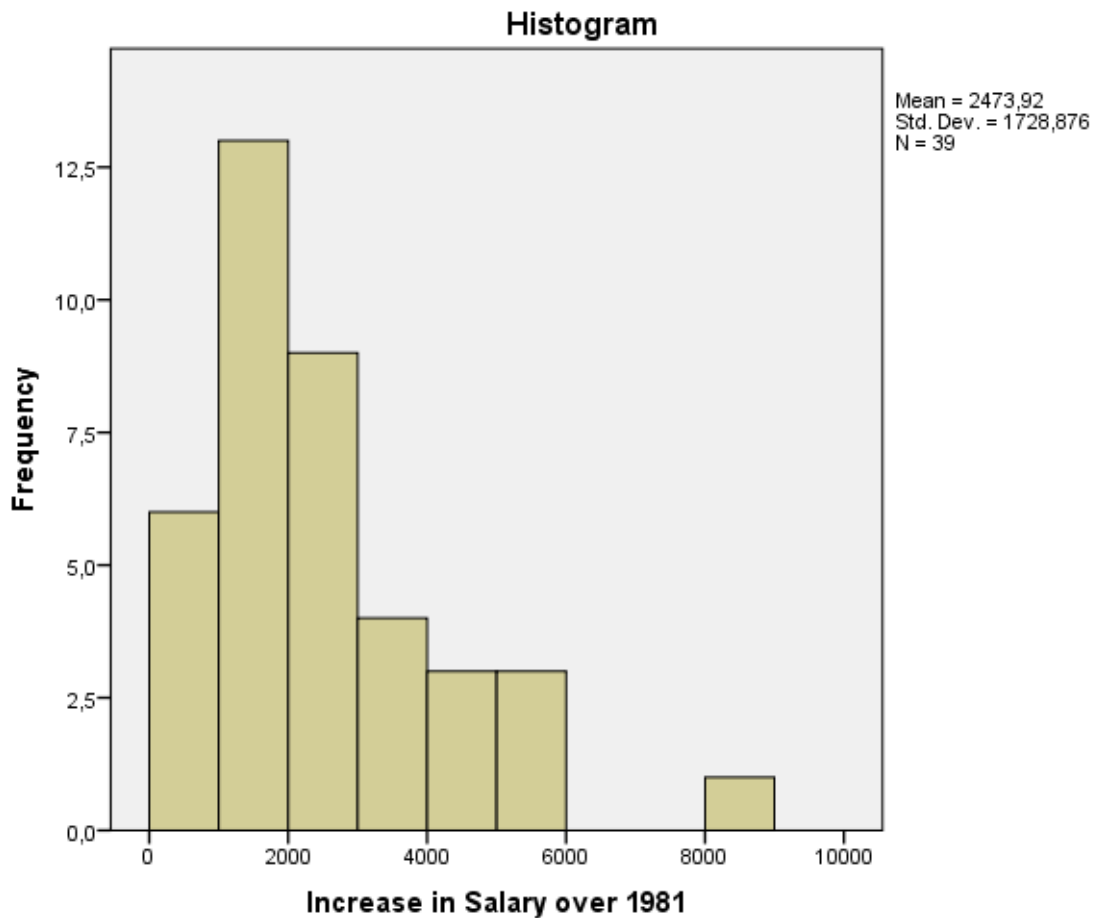


Στο παραπάνω ιστόγραμμα φαίνεται ο τρόπος με τον οποίο κατανέμονται οι ετήσιες απολαβές 40 εργαζομένων της εταιρίας. Συγκεκριμένα χωρίζοντας το εύρος των ετήσιων απολαβών (το εύρος είναι από 9652 μέχρι 50700) σε ισομεγέθη διαστήματα, ταξινομούμε τους εργαζόμενους στο κατάλληλο διάστημα ανάλογα με τις ετήσιες απολαβές του. Ο οριζόντιος άξονας δηλαδή αποτελείται από τα διαστήματα των ετήσιων απολαβών ενώ ο κάθετος άξονας από την συχνότητα, δηλαδή για κάθε διάστημα πόσοι εργαζόμενοι ανήκουν στο διάστημα αυτό. Συγκεκριμένα για το έτος 1982 οι ετήσιες απολαβές των εργαζομένων της εταιρίας διαμορφώνονται ως εξής :

- ✓ 10 εργαζόμενοι έλαβαν 9.652 μέχρι 13.383 ευρώ
- ✓ 8 εργαζόμενοι έλαβαν 13.384 μέχρι 17.114 ευρώ
- ✓ 10 εργαζόμενοι έλαβαν 17.115 μέχρι 20.845 ευρώ
- ✓ 1 εργαζόμενος έλαβαν 20.846 μέχρι 24.576
- ✓ 6 εργαζόμενοι έλαβαν 24.577 μέχρι 28.307

- ✓ 1 εργαζόμενος έλαβαν 28.308 μέχρι 32.038
- ✓ 2 εργαζόμενοι έλαβαν 46.963 μέχρι 50.700

Μεταβλητή « Increase in Salary over 1981 »



Στο παραπάνω ιστόγραμμα φαίνεται ο τρόπος με τον οποίο κατανέμονται οι αυξήσεις στις ετήσιες απολαβές 39 εργαζομένων της εταιρίας. Συγκεκριμένα χωρίζοντας το εύρος των αυξήσεων (το εύρος είναι από 0 μέχρι 8127) σε ισομεγέθη διαστήματα, ταξινομούμε τους εργαζόμενους στο κατάλληλο διάστημα ανάλογα με την αύξηση που έλαβαν στις ετήσιες απολαβές του. Ο οριζόντιος άξονας δηλαδή αποτελείται από τα διαστήματα των αυξήσεων ενώ ο κάθετος άξονας από την συχνότητα, δηλαδή για κάθε διάστημα πόσοι εργαζόμενοι ανήκουν στο διάστημα αυτό. Συγκεκριμένα για το έτος 1982 οι αυξήσεις στις ετήσιες απολαβές των εργαζομένων τις εταιρίας διαμορφώνονται ως εξής :

- ✓ 6 εργαζόμενοι έλαβαν αύξηση από 0 μέχρι 903 ευρώ
- ✓ 13 εργαζόμενοι έλαβαν αύξηση από 904 μέχρι 1.806 ευρώ
- ✓ 9 εργαζόμενοι έλαβαν αύξηση από 1807 μέχρι 2.709 ευρώ
- ✓ 4 εργαζόμενοι έλαβαν αύξηση από 2.710 μέχρι 3.612 ευρώ
- ✓ 3 εργαζόμενοι έλαβαν αύξηση από 3.613 μέχρι 4.515 ευρώ
- ✓ 3 εργαζόμενοι έλαβαν αύξηση από 4.516 μέχρι 5.418 ευρώ
- ✓ 1 εργαζόμενος έλαβε αύξηση από 7.225 μέχρι 8.127 ευρώ

Ερώτημα 5

Αξιοποιώντας τα Box plot του ερωτήματος 3 βλέπουμε ότι οι ακραίες τιμές της μεταβλητής salary82 είναι οι γραμμές με αρίθμηση 45 και 37. Επίσης με την μέθοδο των Z score προκύπτει περίπου το ίδιο συμπέρασμα, με την σειρά νούμερο 45 να είναι outlier ($3,23 > 3$) και την 37 οριακά να μην είναι ($2,96 < 3$).

37	29231	78	33	White	Male	Chicago	Chicago O...	Profession...	No	48100	8125	2,96547
38	29852	79	35	White	Female	Chicago	Chicago O...	Technicians	Yes	19500	3250	,05248
39	30542	80	23	Oriental	Male	Not Emplo...	Not Emplo...	Profession...	Not Emplo...	0	-999	.
40	32512	79	27	White	Female	St. Louis	St. Louis ...	Office and ...	No	14007	1787	-,50700
41	32791	75	51	Oriental	Female	Chicago	Administra...	Officials & ...	Yes	32500	5200	1,37657
42	34651	79	38	Oriental	Female	Chicago	Project Dir...	Profession...	No	17179	2165	-,18392
43	35141	78	33	White	Male	Chicago	Administra...	Officials & ...	No	26000	3250	,71453
44	35461	79	29	White	Female	Chicago	Project Dir...	Profession...	Not Emplo...	15210	-999	-,38447
45	35601	48	61	White	Male	Chicago	Chicago O...	Technicians	No	50700	2925	3,23029

Αντίστοιχα για την μεταβλητή raise82 βλέπουμε την γραμμή με αρίθμηση 37 ως outlier, το ίδιο υποδεικνύει και η μέθοδος με το z score

37	29231	78	33	White	Male	Chicago	Chicago O...	Profession...	No	48100	8125	2,96547	3,26864
----	-------	----	----	-------	------	---------	--------------	---------------	----	-------	------	---------	---------

Ερώτημα 6

Case Summaries

				Employee's Name
Department Code in 1982	Administrative	City Where Employed	Chicago	1 Reuben D. Cross
				2 Brian M. Curtis
				3 Billy J. Lyle
				4 Bernard K. Golden
				5 Helen K. Washington
				6 Regina K. Newton
				7 Rebecca K. Cohen
				8 Charles P. Black
				9 Geraldine K. Jordan
				10 Ping Huffman
				11 John C. Holcomb
				12 Ruby Blair
			Total N	12
			Total N	12
	Project Directors	City Where Employed	St. Louis	1 Peter D. Gamino
				2 Sean D. McGillicuddy
				Total N 2
				1 Jerri D. Mategrano
				2 Amanda Howell
				3 Aristotle M. Geocaris
			Chicago	4 Paulette Ho
				5 Jennifer K. Logan
				6 Pierre LeClair
			Total N	6
			Total N	8
	Chicago Operations	City Where Employed	Chicago	1 Lucinda Jackson
				2 Anita Pulaski
				3 Guy J. LaVelle
				4 Jean D. Weiss
			Chicago	5 Della D. Miller
				6 Nick Kozak
				7 Leonard D. Cunningham
				8 Pamela Simon

				9	Paul Harman	
				10	Andrea Smith	
			Total	N		10
		Total	N			10
				1	Maria Provenza	
				2	Cleveland Smith	
				3	Monica C. Rivers	
				4	Cynthia Riley	
				5	Linda Iverson	
St. Louis	City Where	St. Louis	6	Silvia Kudirka		
Operations	Employed		7	Polly E. Chan		
			8	Kathleen Rooney		
			9	Marci Syms		
			10	Elena Lopez		
			Total	N		10
		Total	N			10
Total	N					40

Στο παραπάνω πίνακα δίνονται τα στοιχεία των υπαλλήλων που εργάστηκαν στο Chicago διοικητική ή σε επιχειρησιακή θέση.

Ερώτημα 7

Case Summaries

						Yearly Salary in 1982
Department Code in 1982	Administrative	City Where Employed	Chicago	1		19240
				2		10241
				3		17111
				4		35750
				5		16120
				6		12558
				7		24784
				8		13910
				9		17108
				10		32500
				11		26000
				12		17810
				Total	N	12
					Mean	20261,00
	Project Directors	City Where Employed	Chicago	N		12
				Total		20261,00
				Mean		17798,50
				1		16016
				2		12566
				3		31200
				4		17179
				5		15210
				6		11700
				Total	N	6
					Mean	17311,83
				N		8
				Total		17433,50
				Mean		9840
	Chicago Operations	City Where Employed	Chicago	1		12328
				2		21060
				3		16601
				4		

			5	9777
			6	11440
			7	48100
			8	19500
			9	50700
			10	9652
			Total N	10
			Mean	20899,80
		Total	N	10
			Mean	20899,80
			1	10780
			2	14309
			3	26182
			4	16900
			5	25577
		St.	6	25577
St. Louis	City Where	Louis	7	11240
Operations	Employed		8	14007
			9	13097
			10	10122
			Total N	10
			Mean	16779,10
		Total	N	10
			Mean	16779,10
Total	N			40
	Mean			18984,73

Yearly Salary in 1982 * Job Category

Yearly Salary in 1982

Job Category	Mean	N	Std. Deviation
Officials & Managers	20261,00	12	7904,138
Professionals	20840,89	9	12341,713
Technicians	17877,56	9	13047,143
Office and Clerical	16779,10	10	6511,245
Total	18984,73	40	9818,086

Ερώτημα : *Να εξεταστεί η διαφορετικότητα και διασπορά δυο περιπτώσεων ανά μεταβλητή.*

Οι τεχνικοί (Technicians) αν και είναι τρίτοι σε κατάταξη στον μέσο μισθό παρουσιάζουν την μεγαλύτερη τυπική απόκλιση στους μισθούς τους. Μπορούμε να υποθέσουμε ότι υπάρχει άνιση κατανομή στους μισθούς τους, το οποίο ίσως οφείλετε στις διάφορες κατηγορίες τεχνικών π.χ ηλεκτρολόγοι, μηχανολόγοι μηχανικοί κτλ .

Οι υπάλληλοι γραφείου (Office and Clerical) παρουσιάζουν πιθανώς μεγαλύτερη ισοκατανομή στα εισοδήματα, καθώς έχουν την μικρότερη τυπική απόκλιση αλλά και μικρότερη μέση τιμή.

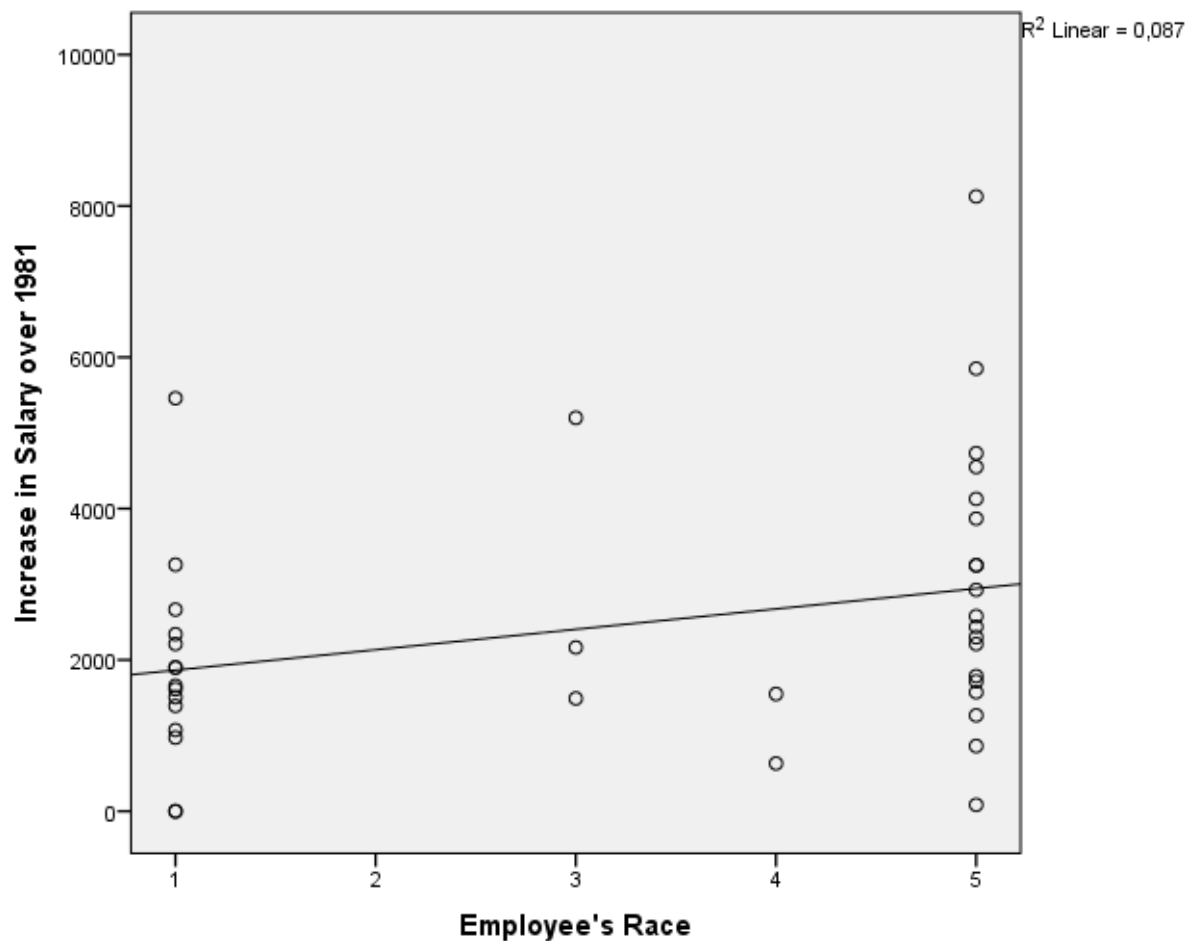
Yearly Salary in 1982 * Department Code in 1982

Yearly Salary in 1982			
Department Code in 1982	Mean	N	Std. Deviation
Administrative	20261,00	12	7904,138
Project Directors	17433,50	8	7393,062
Chicago Operations	20899,80	10	15577,334
St. Louis Operations	16779,10	10	6511,245
Total	18984,73	40	9818,086

Ερώτημα : *Να εξεταστεί η διαφορετικότητα και διασπορά δυο περιπτώσεων ανά μεταβλητή.*

Οι εργαζόμενοι στο τμήμα λειτουργιών του St.Louis παρουσιάζουν την μικρότερη μέση τιμή αλλά και την μικρότερη τυπική απόκλιση, οπότε υπάρχει πιθανώς μια προσεγγιστικά καλύτερη ισοκατανομή μεταξύ των υπαλλήλων σε αυτό το τμήμα, σε αντίθεση με το τμήμα λειτουργιών στο Chicago όπου η τυπική απόκλιση από τον μέσο μισθό είναι τεράστια.

Bonus Ερώτημα



Crosstabs

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Employee's Race * Increase in Salary over 1981	39	78,0%	11	22,0%	50	100,0%

Directional Measures

				Value
Nominal by Interval	Eta	Employee's Race Dependent		,970
		Increase in Salary over 1981 Dependent		,374

Για την εξέταση της συσχέτισης μεταξύ των δύο μεταβλητών χρησιμοποιήσαμε δύο εργαλεία, τον δείκτη eta και ένα scatter plot.

Απο το scatter plot μπρούμε να διακρίνουμε οτι οι αυξήσεις μισθού είναι ξεκάθαρα υψηλότερες όσον αφορά την κατηγορία 5(white race).

Ο Eta παίρνει τιμές από 0 έως 1, οι οποίες έχουν να κάνουν με μηδενική και απολυτή συσχέτιση αντίστοιχα. Στην περίπτωση μας ο Eta έχει τιμή 0,37 που σημαίνει ότι υπάρχει μια μέτρια συσχέτιση μεταξύ αύξησης μισθού και εθνικότητας. Επίσης μπορούμε να τετραγωνίσουμε τον Eta για να δούμε πόση από την διακύμανση ερμηνεύεται από την σχέση των δυο μεταβλητών.