



با سلام و احترام،

فصل هفدهم : دستورالعمل میانگین (Mean)

دستورالعمل **Mean** هم برای توصیف و هم تحلیل متغیرهای کمی (**Scale**) سودمند است. با استفاده از قابلیت‌های توصیفی، می‌توانید از یک مجموعه متنوع آماره‌ها بخواهید که گرایش مرکزی و پراکندگی متغیرهای آزمون را تعیین نمایند. (پراکندگی یا **Dispersion** مشخصه‌ای از توزیع فراوانی است که با پخش مقادیر مرتبط است. کمیت‌های پراکندگی شامل واریانس، انحراف معیار، و دامنه بین چارکی می‌باشد). با استفاده از قابلیت‌های آزمون فرضیات، می‌توانید اختلاف بین میانگین گروه‌ها را به کمک تحلیل واریانس یک طرفه (**ANOVA**) تست نمایید.

استفاده از میانگین‌ها برای کسب آماره‌های توصیفی

به عنوان بخشی از بند قرارداد حقوق پرستاران، یک روزنامه اطلاعات راجع به دستمزد ساعتی آنها را از محل اداره یا بیمارستان به همراه سطح مهارت آنها فراهم می‌کند.

از قابلیت **Means** برای بررسی رابطه بین حقوق، تجربه و محل خدمت استفاده نمایید.

آغاز تحلیل

۱. برای آغاز تحلیل، مسیر **Analyze > Compare Means > Means** را از منوی اصلی برگزینید.
۲. **Hourly Salary** را به عنوان متغیر وابسته برگزینید.
۳. **Years Experience** را به عنوان متغیر مستقل برگزینید.
۴. **Ok** را کلیک کنید.

جدول میانگین با یک متغیر گروه‌بندی

جدول نتایج، آماره‌های پیش فرض را برای حقوق بر حسب تجربه‌های کاری مختلف نمایش می‌دهد. حقوق ساعتی با یک نظم عادلانه در تمام گروه‌های مهارت افزایش می‌یابد.

Years Experience	Mean	N	Std. Deviation
5 or less	18.0416	221	3.86667
6-10	18.9169	460	3.77816
11-15	19.6616	752	3.90528
16-20	20.2876	729	3.82786
21-35	21.2594	539	4.06669
36 or more	21.6342	210	3.61826
Total	20.0159	2911	4.00309

لایه‌بندی متغیرها

حقوق پرستاران صرفاً اصلاً به عنوان یک تابعی از تجربه تغییر نمی‌کند؛ پرستاران بیمارستان و اداره احتمالاً وظایف خود را با توانایی‌های مختلفی شروع می‌کنند. با استفاده از دستورالعمل **Mean**، موقعیت را به همراه سطح مهارت می‌توان لایه‌بندی کرد و نحوه تفاوت حقوق‌ها را بررسی نمود.



۱. کادر محاوره **Means** را بازخوانی کنید.
۲. دکمه **Next** را کلیک کنید تا لایه از 1 به 2 ارتقاء یابد.
۳. **Nurse Type** را به عنوان متغیر مستقل برگزینید.
۴. **Ok** را کلیک کنید.

جدول میانگین لایه‌بندی شده

- * اولین سطر جدول، آماره‌های مربوط به پرستاران بیمارستان را نشان می‌دهد که سابقه پنج سال یا کمتر از پنج سال دارند.
- * سطر دوم همان آماره‌ها را برای پرستاران اداری با همان میزان تجربه نشان می‌دهد.
- * سطر سوم که با عنوان **Total** نشان داده شده است، آماره‌های حاصل از ترکیب دو سطر قبلی را نمایش می‌دهد.

یعنی جمع حاصلضرب میانگین‌ها در تعداد تقسیم بر تعداد کل.

$$[(\text{mean1} * N1) + (\text{mean2} * N2)] / (N1 + N2)$$

✱ در سرتاسر تمام سطرهای جدول، ستون **Mean** نشان می‌دهد که چگونه میانگین حقوق با دو فاکتور محل خدمت پرستاری و میزان تجربه تفاوت می‌کند.

Years Experience	Nurse Type	Mean	N	Std. Deviation
0-unless	Hospital	190753	147	3.77729
	Office	159682	74	3.98762
	Total	180476	221	3.86667
6-10	Hospital	194586	373	3.35288
	Office	177082	147	4.32487
	Total	189469	480	3.79616
11-15	Hospital	202412	518	3.41065
	Office	183784	234	4.17662
	Total	196616	752	3.90528
16-20	Hospital	211099	471	3.29487
	Office	187323	350	4.11943
	Total	202076	729	3.62786
21-25	Hospital	218851	350	3.48089
	Office	204671	185	4.82372
	Total	212504	538	4.06609
26 or more	Hospital	220641	126	3.14466
	Office	206534	84	4.31931
	Total	216342	210	3.61226
Total	Hospital	206764	943	3.49582
	Office	186859	666	4.58552
	Total	200159	2011	4.09099

تحلیل واریانس یک طرفه ANOVA و آزمون خطی

یک پژوهشگر سلامت عمومی، روی اعتیاد به سیگار در بین جوانان مطالعه می‌کند. داده‌ها نشان می‌دهد کسانی که سیگار را از جوانی آغاز کرده‌اند سیگاری‌تر (یعنی وابستگی بیشتر به سیگار) از کسانی هستند که سیگار را از جوانی شروع نکرده‌اند و جالب است بدانید که رابطه خطی می‌باشد. مرحله اول ارزیابی داده‌ها برای انعکاس گرایش‌های جمعیتی خواهد بود.

شروع تحلیل

۱. برای تنظیم وزنی، از منوی اصلی **Data Editor**، مسیر زیر را برگزینید:

Data > Weight Cases

۲. گزینه **Weight Cases by** را انتخاب کنید.

۳. **Final Person - Level Sample Weight** را به عنوان متغیر وزنی برگزینید.

۴. **Ok** را کلیک کنید.

۵. اکنون داده‌ها وزن شده‌اند و آماده تحلیل هستند. برای آغاز تحلیل میانگین، مسیر زیر را از منوی اصلی انتخاب نمایید.

Analyze > Compare Means > Means

۶. گزینه **Age when first smoked a cigarette** (سن شروع اولین سیگار) را عنوان متغیر وابسته برگزینید.

۷. گزینه **#Cigarettes smoked per day past 30 days** سیگار مصرفی در روز بعد از ۳۰ روز سیگار کشیدن را به عنوان متغیر مستقل انتخاب کنید.

۸. دکمه **Options** را کلیک کنید.


۹. گزینه‌های **Anova table and eta** و **Test for Linearity** را فعال کنید. با فعال کردن این دو گزینه جداولی حاوی تحلیل **ANOVA** و ضریب همبستگی **Eta** و میزان خطی بودن خروجی ارائه می‌شود.

۱۰. **Continue** را کلیک کنید.

۱۱. **Ok** را در کادر محاوره **Means** کلیک کنید.

ارتباط میانگین سن با میزان سیگاری بودن

نوجوانانی که یک بسته یا کمتر از یک بسته سیگار برایشان گزارش شده است معمولاً سیگار را از ۱۶ سالگی شروع کرده‌اند (a). در مقایسه، نوجوانانی که بیشتر از یک بسته سیگار در روز می‌کشند سیگار را از ۱۴ سالگی آغاز کرده‌اند (b).



# Cigarettes smoked per day past 30 days	Mean	N	Std. Deviation
1 to 5 cigarettes each day	15.81	1119	4.452
6 to 15 cigarettes (about 1/2 pack) each	15.89	1594	4.820
16 to 25 cigarettes (about 1 pack) each	15.63	1604	5.450
26 to 35 cigarettes (about 1 1/2 pk) each	14.18	622	4.086
35 or more cigarettes (about 2 packs) each	14.45	461	4.376
Total	15.48	5400	4.866

ANOVA و آزمون خطی

جدول **ANOVA** شامل آزمون‌های خطی، غیر خطی و ترکیبی از رابطه بین متغیرهای **Age when first smoked a cigarette** و **#Cigarettes smoked per day past 30 days** می‌باشد.

آزمون خطی دارای یک مقدار با سطح معنی‌داری کوچکتر از ۰.۰۵ می‌باشد، (ستون **Sig**) که نشان می‌دهد که یک رابطه خطی بین سن و میزان سیگار کشیدن وجود دارد. آزمون انحراف از خطی بودن نیز معنی‌دار است و

مفهومش این است که یک رابطه غیر خطی به همراه رابطه خطی وجود دارد (a). این نتایج با آنچه که در جدول Means مشاهده کردید هماهنگ می باشد، سن با افزایش سیگار کشیدن کاهش می یابد اما میزان کاهش سن به مقدار زیادی با اندازه سیگار کشیدن در تضاد است.

			Sum of Squares	df	Mean Square	F	Sig.
Age when first smoked a cigarette *	Between Groups	(Combined)	1974.095	4	493.524	21.158	.000
		Linearity	321.500	1	321.500	56.655	.000
# Cigarettes smoked per day past 30 days		Deviation from Linearity	652.595	3	217.532	9.326	.000
	Within Groups		125841.1	5395	23.326		
	Total		127815.2	5399			

اندازه های پیوستگی

در این داده ها، مربع اندازه های پیوستگی، نزدیک صفر می باشد (ستون Eta Squared) مقدار ضریب نسبت همبستگی در اینجا Ois است و به این معنی می باشد که حدود ۱.۵ درصد سیگار مصرفی در روز توسط سن شروع سیگار تبیین می شود. مقدار انحراف در سنی که شخص شروع به کشیدن سیگار می کند از نظر آماری با معنی دار است اما به طور نسبی کوچک است.

	R	R Squared	Eta	Eta Squared
Age when first smoked a cigarette *				
# Cigarettes smoked per day past 30 days	-.102	.010	.124	.015

خلاصه

با کمک دستورالعمل Means می فهمیم، رابطه معنی دار آماری بین تعداد سیگارهایی که شخص در روز می کشد و سنی که شروع به سیگار کشیدن کرده است وجود دارد.

خروجی های میانگین

گزارش میانگین ها

جدول گزارش، آماره های انتخاب شده را برای هر سطحی از متغیر مستقل فهرست می نماید. آماره ها توزیع متغیر وابسته هر گروه را تبیین می نمایند. مثلاً، برای مناطق غربی (West) میانگین و میانه تقریباً برابر هستند، که حکایت از یک توزیع فروش متقارن دارند. توزیع مناطق مرکزی و شرقی نیز متقارن می باشد (ردیف های Central و East). آماره های چولگی کمتر از 1 نیز نشان دهنده توزیع متقارن می باشد (a). با مقایسه آماره ها در سرتاسر

گروه‌ها، اختلافات بین سطوح متغیر مستقل آشکار می‌شود. مثلاً میانگین فروش در **1993** با حرکت از غرب به شرق افزایش می‌یابد. با این وجود، انحراف معیار تقریباً ثابت مانده است.



		1993 Sales	1996 Sales	
West	Mean	\$182,007.99	\$354,325.54	
	Median	\$182,637.91	\$353,732.09	
	Std. Deviation	\$28,994.54	\$51,003.92	
	Skewness	.01	.27	
Central	Mean	\$187,188.54	\$313,619.53	
	Median	\$189,397.80	\$308,806.96	
	Std. Deviation	\$31,406.39	\$61,836.73	
	Skewness	.60	.18	
East	Mean	\$189,289.20	\$396,184.00	
	Median	\$190,493.77	\$359,660.51	
	Std. Deviation	\$28,776.83	\$62,330.45	
	Skewness	-.12	.01	

جدول ANOVA میانگین‌ها

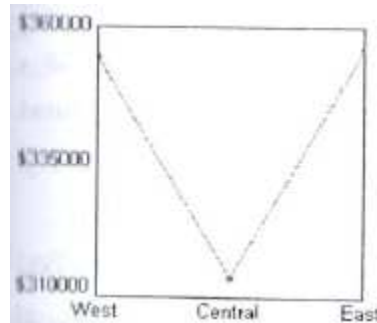
ANOVA، میانگین‌های گروه‌های مختلف را مقایسه می‌کند. تغییر کل به دو مولفه تقسیم شده است. گروه‌های **Between** تغییر میانگین‌های گروه را نسبت به میانگین کلی نشان می‌دهد. مؤلفه **Within Groups** تغییر در امتیازات خاصی را نسبت به میانگین‌های گروهی خودشان نمایش می‌دهد. سطح معنی‌داری کوچک (کمتر از **0.05**) تفاوت‌های گروه را مشخص می‌نماید.

در این مثال، سطح معنی‌داری فروش **1993** از **0.05** تجاوز کرده است، که نشان می‌دهد سه منطقه با هم متفاوت نیستند. همین مقدار برای فروش سال **1996** کمتر از **0.05** است، که نشان می‌دهد حداقل یک منطقه با دیگران تفاوت دارد.

Linearity یک رابطه خطی بین متغیرها را منعکس می‌کند. سطح معنی‌داری کوچک (**0.05**) نشان می‌دهد که رابطه خطی وجود دارد. در این مثال، سطح معنی‌داری از **0.05** تجاوز کرده است (**0.68**). بنابراین هیچ رابطه خطی بین منطقه و فروش **1996** وجود ندارد.

Deviation From Linearity (انحراف از خطی بودن) غیر خطی بودن روابط بین متغیرها را نشان می‌دهد. سطح معنی‌داری کوچک نشان می‌دهند که روابط غیر خطی موجود است. در این مثال، سطح معنی‌داری کوچک‌تر از **0.05** می‌باشد (**0.000**). در نتیجه رابطه غیر خطی بین ناحیه و فروش **1996** وجود دارد. نمودار میانگین‌های این گروه‌ها این روابط را نشان می‌دهد.

		Sum of Squares	df	Mean Square	F	Sig.
1996 Sales * REGION	Between Groups	1.15E+11	2	5.75E+10	16.63	.00
	Linearity	5.87E+09	1	5.87E+09	1.7	.69
	Quadratic from Linearity	1.14E+11	1	1.14E+11	33.18	.00
	Within Groups	1.05E+12	305	3.45E+09		
Total		1.17E+12	307			



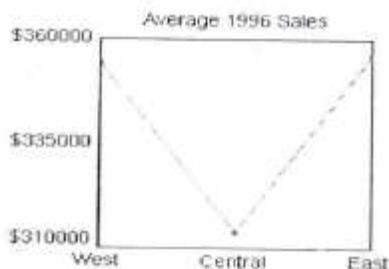
شاخص‌های رابطه در میانگین‌ها

میانگین‌ها چهار شاخص ارتباط بین متغیرها را گزارش می‌کنند. ستون‌های **R** و **R Squared** در صورتی که یک رابطه خطی وجود داشته باشد مناسب هستند. **R Squared** نسبت تغییر در متغیر وابسته را که با مدل خطی در نظر گرفته شده است منعکس می‌کند. در این مثال، یک رابطه خطی با ناحیه، برای **1%** تغییر در فروش **1993** در نظر گرفته شده است (a).

در ستون‌های **Eta** و **Eta Squared** فرض بر این نیست که رابطه‌ای خطی بین متغیرها وجود دارد. **Eta Squared** نسبت تغییر در نظر گرفته شده با تفاوت‌های میان گروه‌ها را نشان می‌دهد. در این مثال، تفاوت‌های بین نواحی برای **10%** تغییر در فروش‌های **1996** لحاظ شده است (b).

	R	R Squared	Eta	Eta Squared
1993 Sales * REGION	.101	.010	.103	b .011
1996 Sales * REGION	.022	a .001	.314	.099

روابط غیر خطی را نمی‌توان دقیقاً با ضریب تعیینی (**R Squared**) توصیف نمود. اختلافات زیاد بین **R Squared** و **Eta Squared** در سال **1996** در اثر روابط غیر خطی بین متغیر حاصل شده است. در صورتی که رابطه غیر خطی باشد مربع **Eta** بر ضرب تعیینی (مربع **R**) ترجیح دارد.



خلاصه‌های موردی

قابلیت **Case Summaries** فهرست‌های موارد خاص و آماره‌های خلاصه را فراهم می‌نماید. فهرست‌های موارد مقادیر ثبت شده مربوط به متغیرهای منتخب هر مورد (مشاهده شده) در فایل داده‌ها می‌باشد. آماره‌های خلاصه بر داده‌های ظاهر شده در جدول مبتنی می‌باشند، نه بر کل فایل داده‌ها.

با خلاصه‌های موردی می‌توان آماره‌های خلاصه مربوط به گروه‌ها را بدون فهرست موارد خاص نیز نمایش داد.

**1996 Sales
by Location and Division**

Region		Division	
		Consumer Products	Business Products
East	N	84	56
	Median	\$273,600	\$314,250
	Mean	\$289,814	\$404,178
	Total	\$18,548,100	\$22,633,950
Central	N	77	84
	Median	\$293,400	\$317,250
	Mean	\$311,774	\$396,883
	Total	\$24,006,600	\$33,338,200
West	N	38	71
	Median	\$306,900	\$399,000
	Mean	\$336,221	\$475,141
	Total	\$12,776,400	\$33,735,000

مبحث را در مرجع کاربردی **SPSS 20** (۲۵) دنبال نمایید.

