



## پروژه تحلیل داده‌های فضایی

گردآورندگان: سارا معصومی - نرگس متقی - محمدعلی قمرانی - دنیا مهدی

استاد راهنما: دکتر محسن محمدزاده

خرداد ۱۴۰۲

## فهرست

صفحه	عنوان
۲	چکیده
۳	مقدمه
۴	جدول داده ها
۵	تحلیل اکتشافی
۱۴	برازش مدل واریوگرام
۱۶	پیشگویی فضایی
۱۷	پیوست
۲۷	منابع

## چکیده

در این مقاله قصد داریم مطالعاتی را بر روی میانگین دمای برخی از شهرهای ایران انجام دهیم. برای این منظور مجموعه داده‌ای را از سایت meteomanz اخذ کردیم. این مجموعه داده شامل موقعیت‌های جغرافیایی ۷۷ شهر مختلف از ایران می‌باشد که این موقعیت‌ها بر اساس طول و عرض جغرافیایی مشخص شده‌اند. این مجموعه داده علاوه بر موقعیت جغرافیایی شامل اطلاعات ثبت شده‌ای از میانگین دمای این شهرها در تاریخ ۱۴۰۲/۰۴/۰۹ می‌باشد.

با توجه به این که داده‌های ما از لحاظ موقعیت مکانی به یکدیگر وابسته هستند نمی‌توان از روش‌های مرسوم آمار کلاسیک برای بررسی آنها استفاده کرد چرا که در آمار کلاسیک برقراری فرض استقلال مشاهدات، ضروری است. بنابراین برای اهداف مختلف از جمله پیشگویی، از روش‌های نوین آماری مانند روش‌های تحلیل داده‌های فضایی بهره می‌بریم.

لازم به ذکر است برای اجرای روش‌های مذکور بر روی داده‌های خود از نرم افزار R کمک می‌گیریم، همچنین کدهای اجرا شده در بخش پیوست مقاله قرار خواهند گرفت.

### آشنایی با تحلیل فضایی

در بررسی‌های محیطی اغلب با داده‌هایی مواجه می‌شویم که مستقل نیستند و به نوعی وابستگی آن‌ها ناشی از موقعیت و مکان قرار گرفتن داده‌ها در فضای مورد بررسی یا زمان مشاهده آن‌ها است. تحلیل این‌گونه مشاهدات که داده‌های فضایی نامیده می‌شوند، از روش‌های معمول آمار ممکن نیست و لازم است به نحوی ساختار همبستگی داده‌ها در تحلیل آماری داده‌ها منظور شود. روش‌های آمار فضایی به دلیل استفاده از همبستگی فضایی بین داده‌ها که عموماً توسط تابع تغییرنگار مدل بندی می‌شود از دقت بالایی نسبت به سایر روش‌های دیگر برخوردار است.

آمار فضایی به عنوان شاخه‌ای نوین از علم آمار که به روش‌هایی برای تعیین ساختار همبستگی و تحلیل داده‌های فضایی و داده‌های فضایی-زمانی می‌پردازد، در نظر گرفته می‌شود.

یکی از موضوعات مهم در تحلیل داده‌های فضایی، پیشگویی مقدار نامعلوم یک میدان تصادفی با موقعیتهای مشخص بر اساس بردار مشاهدات است. که به آن روش پیشگویی کریگیدن می‌گویند. در روش‌های کریگیدن واریانس پیش بینی در هر نقطه نیز ارائه می‌شود که این یکی از ویژگیهای منحصر به فرد روش کریگیدن است. اگر میدان تصافی مورد نظر گاوسی با ساختار میانگین و کوواریانس پارامتری باشد. پیشگویی بهینه و میانگین مجذور خطای آن قابل محاسبه است. اما در عمل با موارد زیادی مواجه می‌شویم که مشاهدات از مدل گاوسی تبعیت نمی‌کنند. در اینگونه موارد اگر تبدیلی غیر خطی از میدان تصادفی، گاوسی باشد پیشگویی فضایی امکانپذیر می‌گردد. اما اگر تبدیل نامعلوم باشد، می‌توان آن را متعلق به یک خانواده از تبدیلات پارامتری دانست. در اینصورت پیشگویی فضایی علاوه بر پارامترهای مدل به پارامتر تبدیل نیز وابسته خواهد شد و معمولاً برآوردهای حداکثر درست نمایی آنها تعیین و در پیشگوی بهینه جایگذاری می‌شود. اما این کار موجب تردید در بهینگی پیشگویی حاصل خواهد شد. از طرف دیگر، معمولاً تعیین میانگین مجذور خطای این پیشگو یا حتی ارائه تقریب مناسبی برای آن بسیار دشوار یا نشدنی است. لذا برای فائق آمدن بر این مسئله، رهیافت بیزی اتخاذ می‌گردد و با فرض آنکه پارامترهای مدل و تبدیل، متغیرهایی تصادفی هستند پیشگوی بهینه و میانگین مجذور آن تعیین می‌شوند.

## جدول داده ها

نام شهر	تاریخ ثبت	عرض lat	طول lon	میانگین دما (°C)	نام شهر	تاریخ ثبت	عرض lat	طول lon	میانگین دما (°C)
ایلام	1402/04/09	33.57505417	46.41641617	26.7	پارس آباد	1402/04/09	39.64579391	47.90995026	27.9
خرم آباد	1402/04/09	33.45696259	48.34891129	27.5	ماکو	1402/04/09	39.2921524	44.49243546	24.5
الیگودرز	1402/04/09	33.40154648	49.69316864	23.7	خوی	1402/04/09	38.52029419	44.95465469	26.4
کاشان	1402/04/09	33.98339462	51.43208313	34.1	اهر	1402/04/09	38.47453308	47.07098389	22.9
خور	1402/04/09	33.77181625	55.08384323	37.5	تبریز	1402/04/09	38.0792923	46.28915024	25.7
طیس	1402/04/09	33.57518768	56.93533325	39.4	اردبیل	1402/04/09	38.24974442	48.29428482	18.7
فردوس	1402/04/09	34.01807785	58.17238617	33.2	سراب	1402/04/09	37.94099808	47.53533173	21.5
دزفول	1402/04/09	32.37889481	48.41930008	35.1	ارومیه	1402/04/09	37.53953934	45.05765152	23.2
شهرکرد	1402/04/09	32.31618118	50.85820389	21.9	مراغه	1402/04/09	37.38603592	46.23360443	26.6
اصفهان	1402/04/09	32.57763672	51.6590538	33.2	میانه	1402/04/09	37.4207077	47.71635056	29.9
بیرجند	1402/04/09	32.85693359	59.21954727	30.7	بندر انزلی	1402/04/09	37.46670532	49.46269989	27.6
اهواز	1402/04/09	31.24201393	48.67127228	38.6	رشت	1402/04/09	37.28414536	49.59040833	25.6
مسجد سلیمان	1402/04/09	31.96466255	49.28633881	36.6	مراوه تپه	1402/04/09	37.90470886	55.96172714	28.0
آباده	1402/04/09	31.14298248	52.65011597	26.9	بروجرد	1402/04/09	33.89829254	48.75468063	27.2
یزد	1402/04/09	31.88352203	54.34774017	35.5	مهاباد	1402/04/09	33.52765656	52.21732712	26.1
نهبندان	1402/04/09	31.54078674	60.03780746	33.8	سقز	1402/04/09	36.24411774	46.27292252	21.1
زابل	1402/04/09	31.00886726	61.49380875	39.5	زنجان	1402/04/09	36.67955399	48.49299622	22.6
آبادان	1402/04/09	30.35600853	48.28005219	37.4	قزوین	1402/04/09	36.28174591	50.00196838	26.3
امیدیه	1402/04/09	30.7556076	49.71224976	37.5	رامسر	1402/04/09	36.91598892	50.66233444	26.2
گچساران	1402/04/09	30.3659761	50.7936301	31.4	نوشهر	1402/04/09	36.64708328	51.4998436	25.7
یاسوج	1402/04/09	30.65676308	51.58242416	24.6	بابلسر	1402/04/09	36.69893646	52.64933777	26.4
کرمان	1402/04/09	30.28140259	57.06418228	30.1	قراخیل	1402/04/09	36.4532673	52.7830627	25.1
خارک	1402/04/09	29.23805237	50.31930542	33.5	گرگان	1402/04/09	36.84519958	54.42879868	26.4
شیراز	1402/04/09	29.55689049	52.5291214	31.3	شاهرود	1402/04/09	36.41208649	54.9724617	29.5
سیرجان	1402/04/09	29.43761635	55.67490768	33.2	قوچان	1402/04/09	37.1023674	58.51087189	26.1
بافت	1402/04/09	29.23217583	56.59699249	27.5	سرخس	1402/04/09	36.54865646	61.14860153	34.0
بم	1402/04/09	29.06779861	58.34861755	38.9	سبزوار	1402/04/09	36.20921326	57.68231964	34.9
زاهدان	1402/04/09	29.48213577	60.85205078	31.1	مشهد	1402/04/09	36.31043243	59.57567215	30.5
بوشهر	1402/04/09	28.92142677	50.83766937	32.1	سنندج	1402/04/09	35.24031448	47.00640869	24.3
فسا	1402/04/09	28.91972542	53.6376915	30.4	تهران	1402/04/09	35.70290756	51.34975815	32.4
بندر دیر	1402/04/09	27.83905983	51.93501282	32.6	سمنان	1402/04/09	35.57711411	53.38192749	33.7
بندر عباس	1402/04/09	27.19265175	56.29146576	35.0	تربت حیدریه	1402/04/09	35.27845764	59.21302032	31.0
کهنوج	1402/04/09	27.92245483	57.69707489	36.4	کاشمر	1402/04/09	35.2427597	58.45988083	32.8
سراوان	1402/04/09	27.36134338	62.33307266	34.5	کرمانشاه	1402/04/09	34.31409454	47.09714127	29.1
ایران شهر	1402/04/09	27.20416451	60.67512894	38.9	همدان	1402/04/09	34.7918129	48.52614594	23.1
کیش	1402/04/09	26.53342438	53.97384644	32.4	اراک	1402/04/09	34.09344864	49.7220993	27.4
بندر لنگه	1402/04/09	26.54753304	54.88892365	33.1	بندر جاسک	1402/04/09	25.65699577	57.80063248	32.0
سیری	1402/04/09	26.54753304	54.88892365	33.5	چاه بهار	1402/04/09	25.29451942	۶۰,۶۴۷۷۰۱۲۶	31.8
ابوموسی	1402/04/09	25.88346291	55.02985382	33.3					

## تحلیل اکتشافی

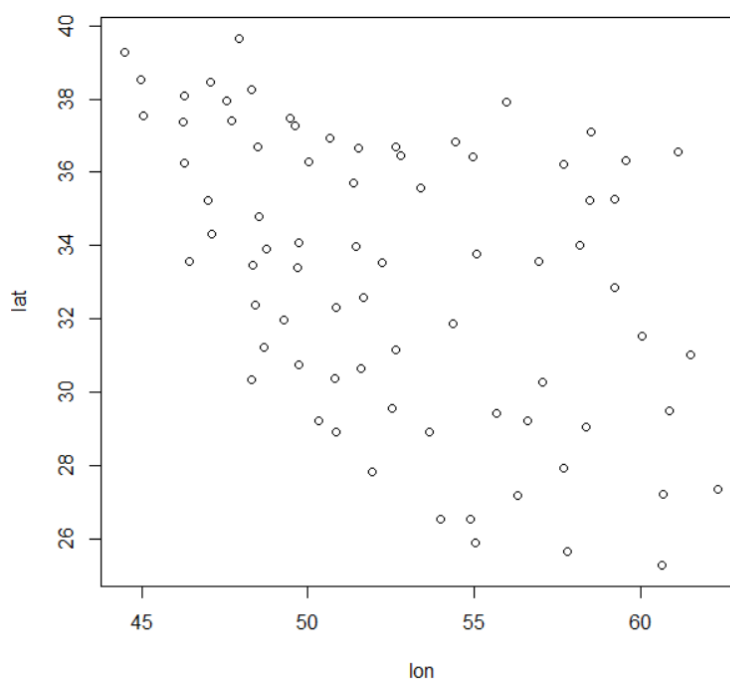
۱. پس از فراخوانی داده‌ها در نرم افزار، اولین قدم بررسی اجمالی متغیر مورد مطالعه می‌باشد. که جدولی از خلاصه های آماری آن را مشاهده و بررسی میکنیم:

خلاصه آماری میانگین دمای شهرها

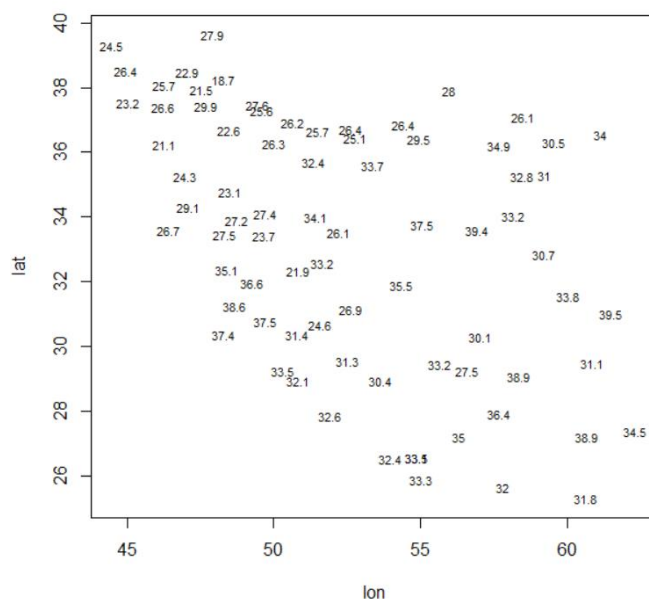
مقادیر	عنوان
۱۸.۷	کمترین دمای ثبت شده
۲۶.۳	چارک اول
۳۰.۵	میانه
۳۳.۵	چارک سوم
۳۰.۰۹	میانگین
۳۹.۵	بیشترین دمای ثبت شده

همانطور که در جدول فوق میبینیم سردترین نقطه از میان این شهرها، اردبیل با دمای ۱۸.۷ درجه‌ی سلسیوس و گرم‌ترین نقطه، شهر زابل با دمای ۳۹.۵ درجه سلسیوس است. متوسط دمای تمام شهرها تقریباً ۳۰ درجه سلسیوس میباشد همچنین ۵۰٪ شهرها در ۹ تیرماه ۱۴۰۲ دمای کمتر از ۳۰.۵ و ۵۰٪ آن‌ها دمایی بالاتر از مقدار مذکور را تجربه کرده اند.

به منظور اینکه پراکندگی داده‌های خود را در سطح کشور به طور شهودی ببینیم نمودار زیر را رسم میکنیم:



همانطور که در نمودار صفحه قبل مشاهده می‌کنیم در جهت شمال ، شمال غرب و غرب اطلاعات بیشتری نسبت به شرق ، مرکز و جنوب شرقی داریم. میتوانیم برای درک بهتر و مقایسه دما در جهات مختلف مقدار دما را نیز برای هر نقطه بر روی نمودار نمایش دهیم.



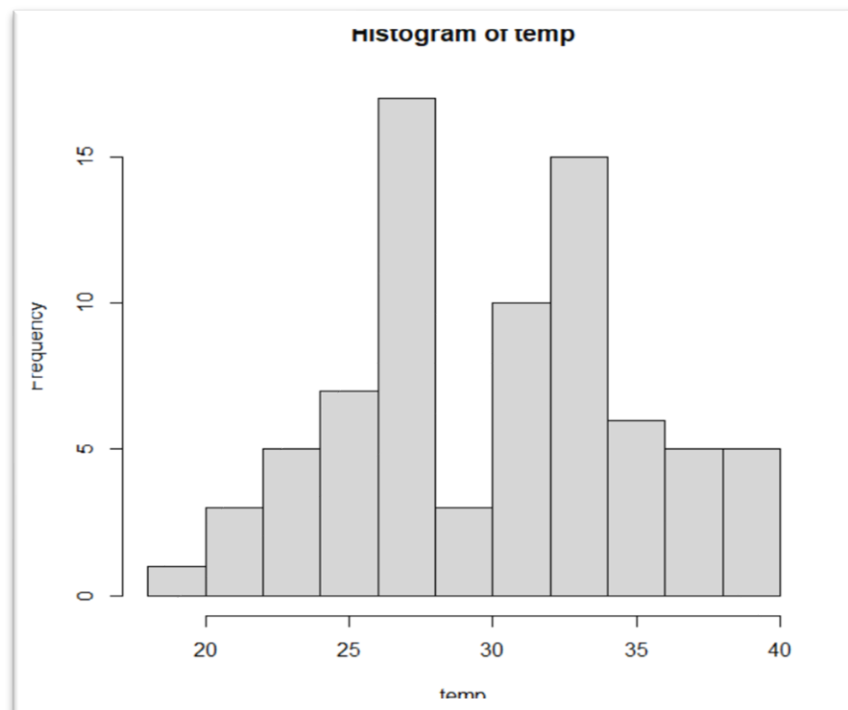
با توجه به نمودار فوق دمای شهرهای شمالی ، غربی و شمال غربی به دلیل وجود رشته کوه‌های البرز و زاگرس، پایینتر از دمای شهرهای جنوبی، شرقی و جنوب شرقی است.

## ۲. بررسی توزیع متغیر مورد مطالعه

به منظور بررسی توزیع یک متغیر، دو راهبرد شهودی و آزمون فرض وجود دارد که نتایج حاصل از آزمون فرض نسبت به رسم نمودار قابل استنادتر است. بنابراین میتوانیم از آزمون شاپیروویلک و رسم نمودار هیستوگرام برای بررسی فرض نرمال بودن توزیع متغیر میانگین دمای شهرها استفاده کنیم :

### آزمون شاپیروویلک

۰.۹۷۷۳۶	مقدار آماره آزمون
۰.۱۸۵۳	P-value



با توجه به میزان  $p$ -value برای آزمون شاپیرووولیک و رسم نمودار هیستوگرام، فرض صفر آزمون یعنی فرض نرمال بودن توزیع رد میشود. در نتیجه داده‌های میانگین دمای شهرها از توزیع نرمال نمی‌باشد، بنابراین میدان تصادفی ما گاوسی نیست.

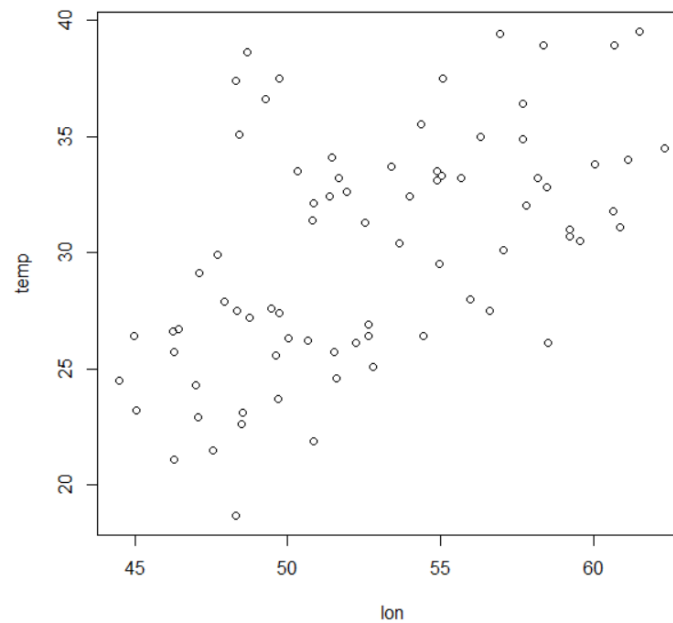
بررسی مانایی در میانگین

برای بررسی وجود یا عدم وجود روند در داده‌ها لازم است مقادیر میدان تصادفی را یک بار در مقابل طول جغرافیایی و یک بار در مقابل عرض جغرافیایی رسم کنیم اگر هیچ الگوی خاصی در پراکندگی داده‌ها مشاهده نشود یعنی میدان تصادفی مانا است در غیر این صورت فرض مانایی رد شده و لازم است داده‌های خود را روند زدایی کنیم.

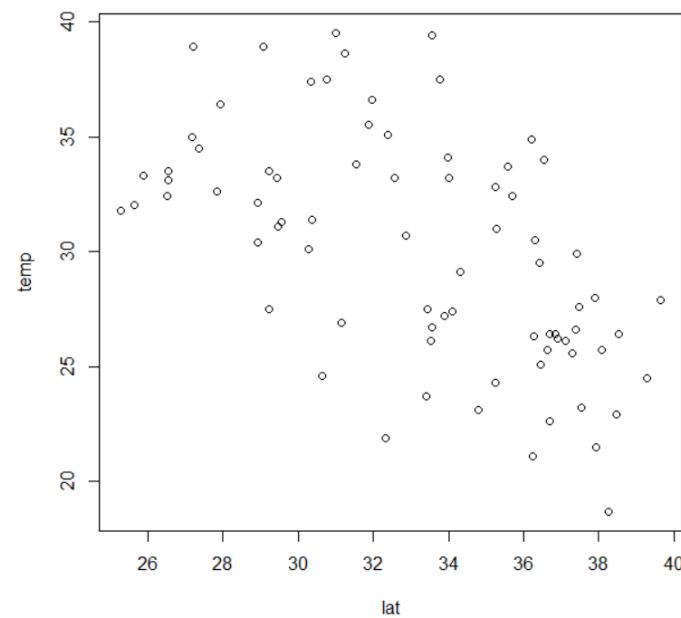
رسم نمودار:



مقادیر میدان تصادفی در مقابل طول جغرافیایی



مقادیر میدان تصادفی در مقابل عرض جغرافیایی



با توجه به نمودارهای رسم شده در فوق برای بررسی مانایی میتوان نتیجه گرفت، داده ها هم در برابر افزایش طول و هم در برابر افزایش عرض، دارای روند هستند. اما در جهت طول جغرافیایی دارای روند صعودی است یعنی هرچه از سمت غرب کشور به سمت شرق کشور حرکت کنیم میانگین دمای هوا بیشتر می شود و در جهت عرض جغرافیایی دارای روند نزولی می باشد یعنی هر چه از سمت جنوب کشور به سمت شمال کشور حرکت کنیم از میانگین دمای هوا کاسته خواهد شد.

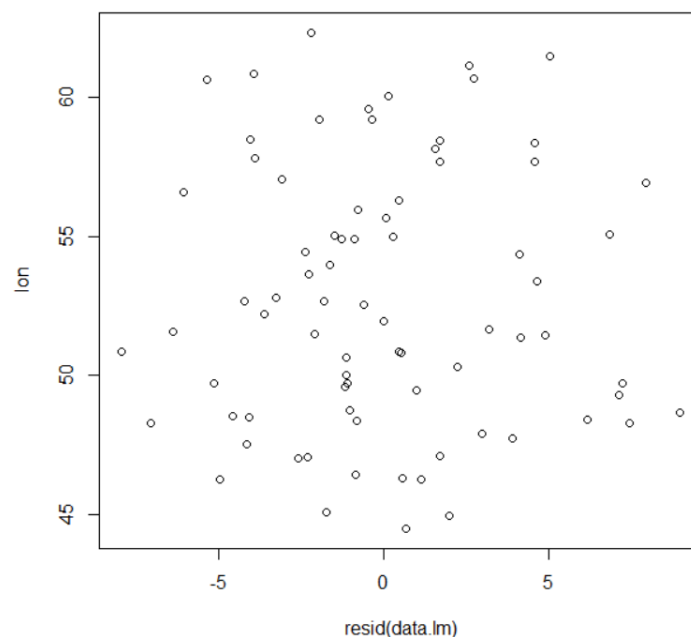
با توجه به نتایج حاصل، لازم است داده‌های خود را روند زدایی کنیم. به کمک رگرسیون خطی مقدار ضریب هر متغیر را محاسبه میکنیم سپس با بدست آوردن اختلاف مقادیر مشاهده شده از مقادیر پیشگویی شده، خطاها را محاسبه میکنیم این خطاها در برابر هیچ کدام از متغیرهای عرض و طول روندی داشته باشند.

### ضرایب رگرسیونی

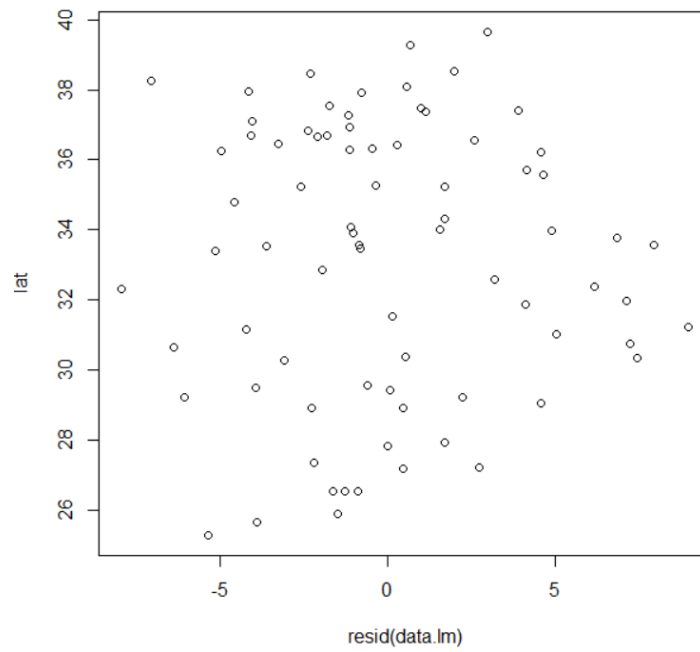
عرض از مبدا	28.12
ضریب lat	0.36
ضریب lon	-0.52

حال میتوانیم مقادیر باقی مانده ها را به کمک نرم افزار R حساب کنیم و یکبار دیگر در برابر عرض و طول جغرافیایی آنها را رسم کنیم همچنین نمودار هیستوگرام را برای بررسی توزیع آنها رسم میکنیم:

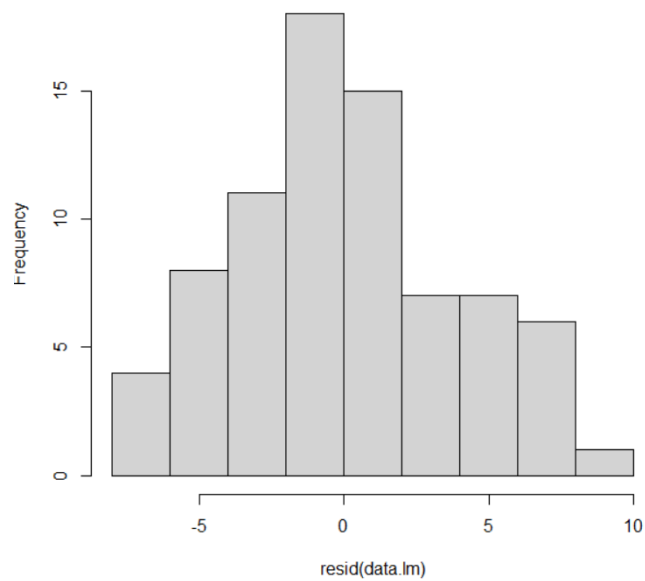
### داده های روند زدوده در برابر طول



داده های روند زدوده در برابر عرض



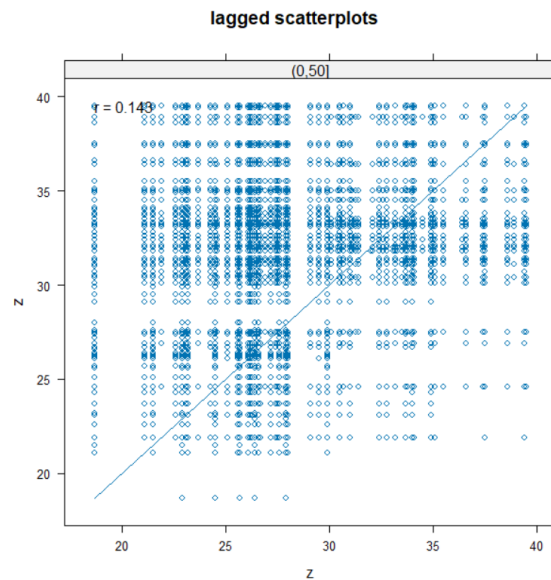
Histogram of resid(data.lm)



همانطور که در نمودار های فوق میبینیم داده ها ما روند زدوده شده اند زیرا نه در برابر طول و نه در برابر عرض دارای روند نیستند و بدون الگو پراکنده شده اند همچنین توزیع آنها شبیه به توزیع نرمال شده است.

## بررسی مانایی در تغییرنگار

برای این کار لازم است نمودار H-پراکنش را برای لگ های مختلف رسم کنیم و اگر نقاط رسم شده نسبت به نیم ساز متقارن باشند یعنی میدان تصادفی مانای در تغییر نگار است.



نمودار H-پراکنش فوق نسبت به خط نیم ساز تقریباً متقارن است بنابراین میدان تصادفی ما مانای در تغییر نگار می باشد.

### ۳. بررسی همسانگردی

برای بررسی همسانگردی لازم است تغییر نگار تجربی را در چهار جهت 0, 45, 90, 135 درجه بر روی نمودار رسم شود اگر تغییر نگار ها در تمام جهات تقریبا بر هم منطبق شده باشند همسانگردی برقرار است

❖ علاوه بر تلاش های گروه ما برای بررسی همسانگردی، کدهای مربوط به این بخش با خطای نامعلوم مواجه میشدند.

### ۴. بررسی فضایی بودن داده ها

برای بررسی فرض فضایی بودن داده های میانگین دمای شهد ها میتوانیم از آزمون موران استفاده کنیم نتایج این آزمون برای متیغر ذکر شده به شرح زیر میباشد:

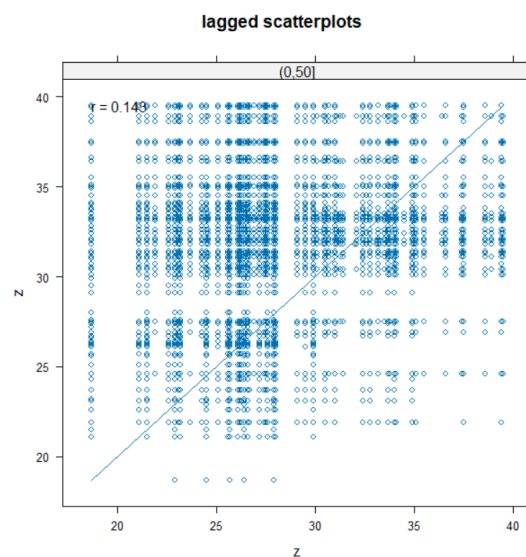
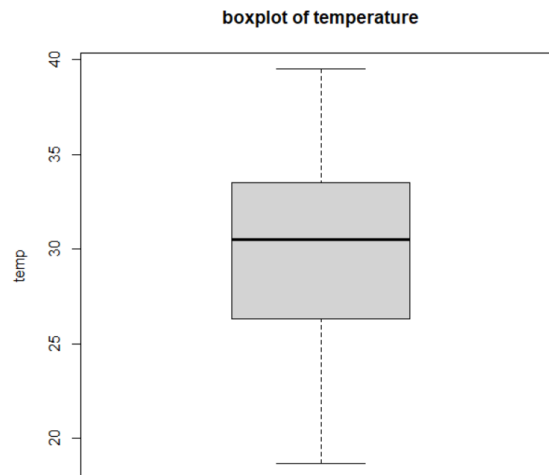
#### آزمون موران

۰.۰۴۱۷	P-value
--------	---------

با توجه به مقدار p-value برای آزمون موران که برابر ۰.۰۴۱۷ است و از ۰.۰۵ کمتر است فرض استقلال داده ها رد میشوند بنابراین داده های ما به یکدیگر وابسته هستند.

## ۵. بررسی داده های پرت

در تحلیل فضایی دو نوع داده ی پرت وجود دارد نوع اول: داده هایی که بدون درنظر گرفتن موقعیت آن ها، فقط براساس مقدار مشاهده شده غیرمعمول هستند، که با استفاده از نمودار جعبه ای به طور شهودی آن ها را شناسایی می کنیم و نوع دوم داده هایی که مقدار آنها در مقایسه با مشاهدات مجاور غیرمعمول هستند. با استفاده از نمودار H-پراکنش می توان این نوع داده ها را شناسایی کرد.



با توجه به نمودار جعبه ای رسم شده در صفحه قبل میتوان نتیجه گرفت مشاهده ی پرت از نوع اول در مجموعه داده ها وجود ندارد. و از طرق نمودار H-پراکنش با توجه به اینکه در گوشه های نمودار چند مشاهده وجود دارد میتوان نتیجه گرفت داده های پرت از نوع دوم در مجموعه داده ی ما وجود دارد.

## برازش مدل واریوگرام

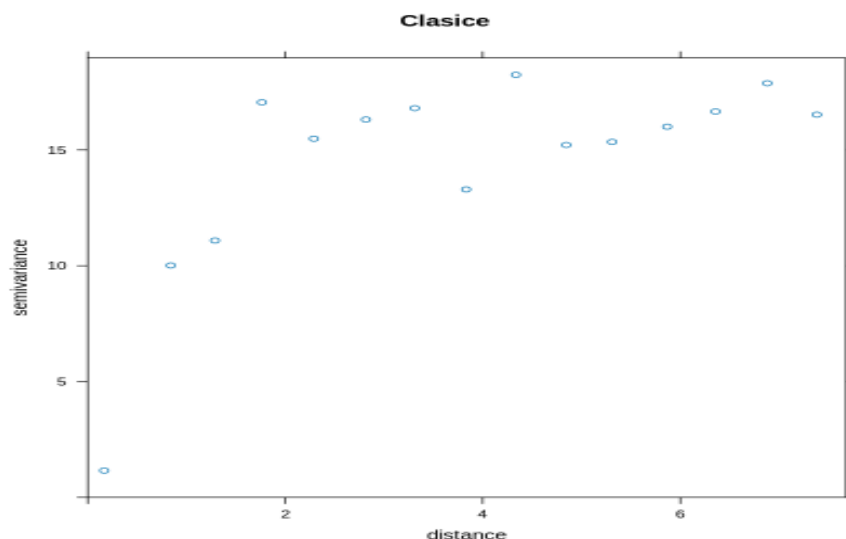
با برآورد تغییرنگار می توانیم درک بهتری از ساختار همبستگی فضایی داده‌ها در آمار فضایی بدست آوریم. و برای محاسبه تغییرنگار از فرمول ذیل استفاده می کنیم

$$2\gamma(s_1, s_2) = \text{Var}(Z(s_1) - Z(s_2))$$

هرچقدر میزان تغییرنگار کوچکتر باشد ، نشانگر وابستگی زیاد و بزرگ بودن آن ، نشانگر وابستگی کم میدان تصادفی است و اگر تغییرنگار برای تمام مقادیر  $h$  تقریباً ثابت باشد ، داده‌ها همبستگی فضایی ندارند یا همبستگی آن‌ها بسیار ضعیف است . و برعکس، به طور کلی می توان گفت کھشیب نمودار تغییرنگار بیانگر همبستگی فضایی داده‌هاست .

برآورد نیم تغییرنگار تجربی کلاسیک و نیرومند:

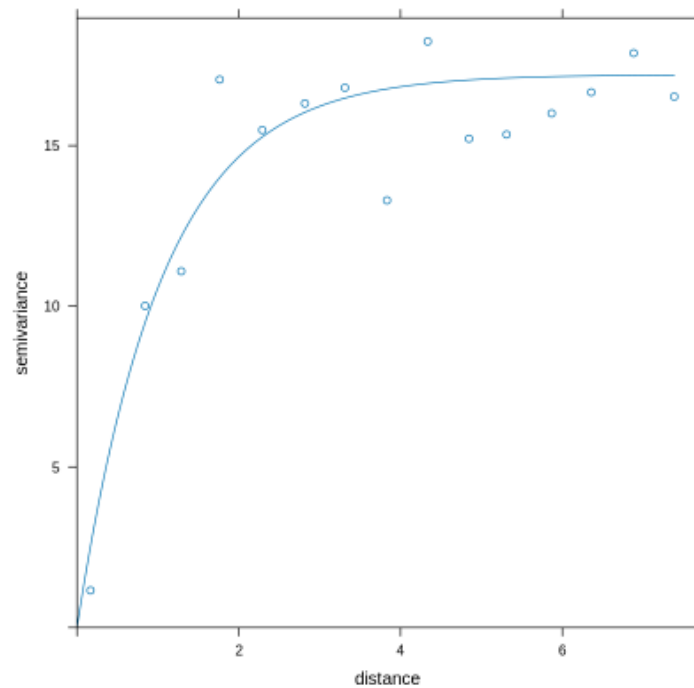
در شکل زیر نمودار نیم تغییرنگار کلاسیک برای داده‌های مورد مطالعه‌ی ما رسم کرده ایم.



نمودار نیم تغییرنگار تجربی کلاسیک

### برازش مدل برای نمودار نیم‌تغییرنگار

باتوجه به نمودار نیم‌تغییرنگار تجربی که رسم کردیم مدل نمایی را به عنوان یک مدل تغییرنگار مناسب انتخاب کرده ایم و آن را بروی داده ها برازش داده ایم که نمودار آن را در شکل ذیل مشاهده می کنید :





## پیشگویی فضایی

با در نظر گرفتن شرط میدان تصادفی گاوسی پیش‌گو فضایی بهینه ترکیب خطی از سایر مشاهدات است که باتوجه به وضعیت مانایی ( میانگین میدان تصادفی) کریگیدن ، عادی را برای پیشگویی دما در شهرهای دیگر انتخاب کرده ایم.

در بخش تحلیل اکتشافی مشاهده کردیم که میدان تصادفی دارای روند می‌باشد و با استفاده از مدل رگرسیونی به برآورد این روند پرداختیم و داده‌ها را روند زدوده کردیم. داده های روند زدوده به طور شهودی مانای در میانگین است و در سطح ۰.۱. نرمال می‌باشد لذا برای پیش‌گویی از کریگیدن عادی استفاده می‌کنیم.

## کریگیدن عادی

میدان تصادفی  $Z(s) = \mu + \delta(s)$  را در نظر بگیریم که در آن  $\mu \in R$  نامعلوم و  $\delta(s)$  یک میدان تصادفی مانای ذاتی با میانگین صفر باشد. هدف پیش‌گویی فضایی  $Z(s_0)$  براساس مشاهدات  $Z = (Z(s_1), \dots, Z(s_n))$  است.

## پیش‌گویی نقاط جدید

شهر	نقاط	مقدار پیش بینی شده	واریانس
خارک	(29.238, 50.319)	1.7۷	2.39
هفت چشمه	(34.204, 47.759)	-0.4۹	2.38
خوسف	(32.779, 58.885)	0.84	2.5
دماوند	(35.693, 52.047)	0.28	2.4

## پیوست

در این بخش کدهای اجرا شده در نرم افزار R به همراه خروجی های آن قرار دارند.

۱. نحوه ی ساخت دیتا فریم در R

#طول

lon <- c()

#عرض

lat <- c()

#دما

temp <- c()

# ساخت دیتا فریم

data <- data.frame(x=lon, y=lat, z=temp, stringsAsFactors = FALSE)

print(is.data.frame(data))

print(ncol(data))

print(nrow(data))

```
> print(is.data.frame(data))
[1] TRUE
> print(ncol(data))
[1] 3
> print(nrow(data))
[1] 77
> |
```

۲. محاسبه ی خلاصه ی آماری

`summary(data$z)`

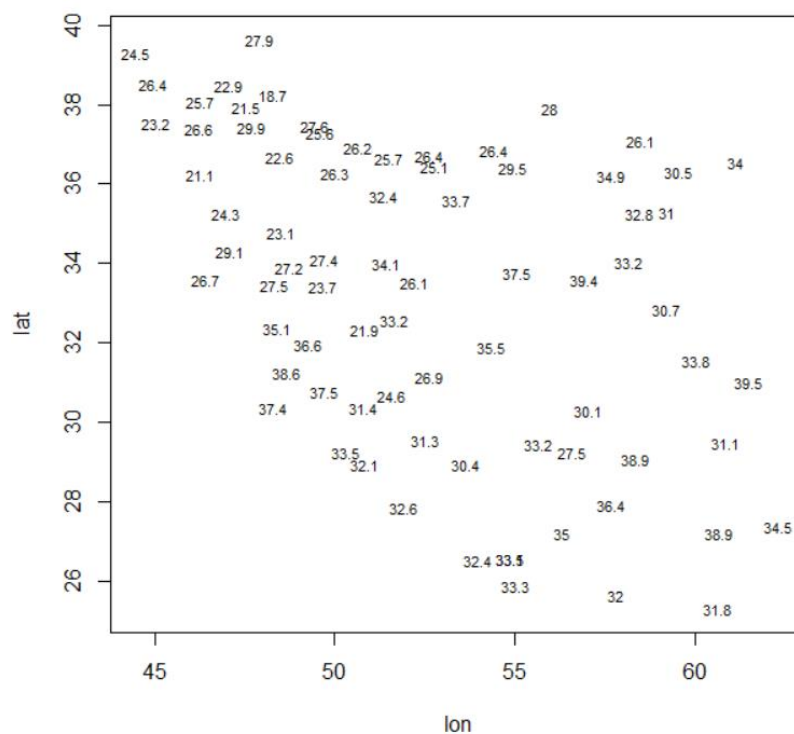
```
> summary(data$z)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  18.70  26.30   30.50   30.09  33.50   39.50
```

۳. رسم نمودار پراکندگی موقعیت ها و نمایش دمای هر موقعیت بر روی نمودار

`plot(lon,lat,type="n")`

`text(lon,lat,temp,offset = 0.5, cex = 0.7)`

`plot(lon,lat)`



۴. آزمون شاپیروویلیک

shapiro.test(temp)

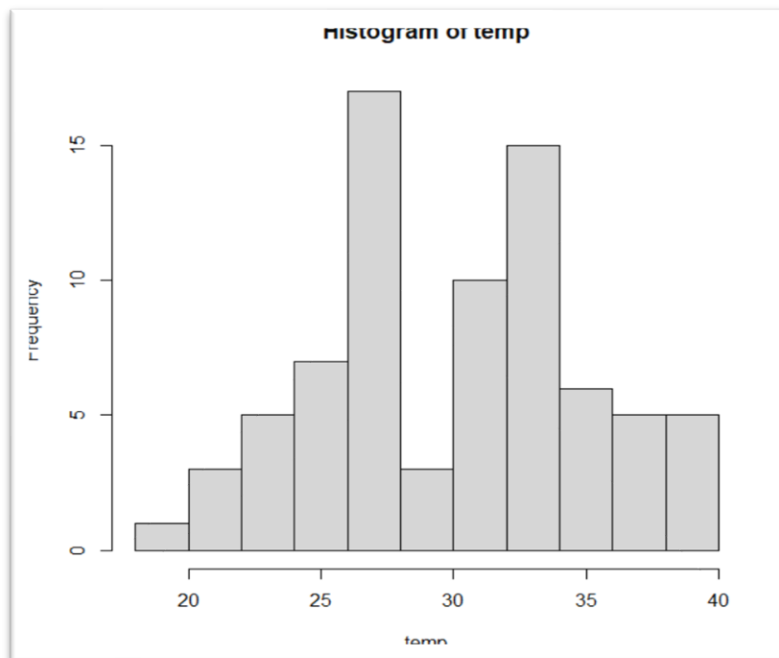
```
> shapiro.test(temp)

      Shapiro-Wilk normality test

data:  temp
W = 0.97736, p-value = 0.1853
```

۵. رسم نمودار هیستوگرام

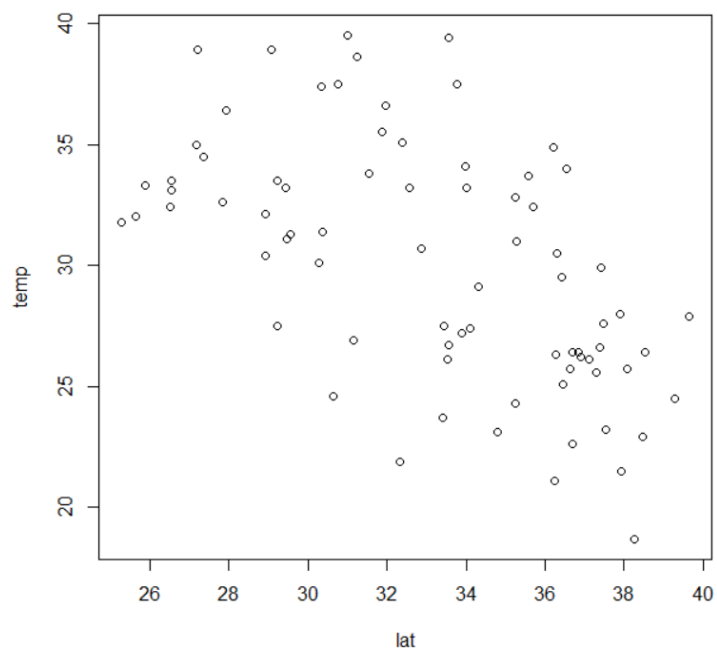
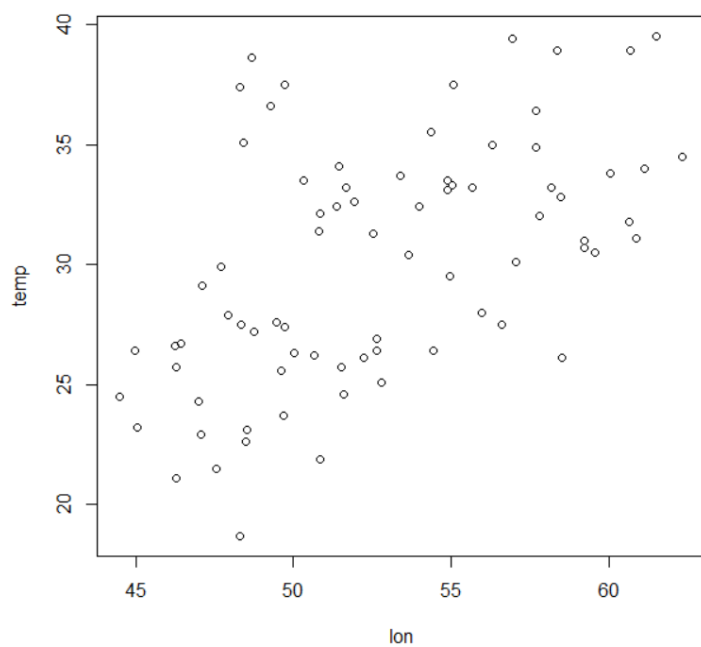
hist(temp)



۶. رسم نمودار پراکنش جهت بررسی مانایی

`plot(lon,temp)`

`plot(lat,temp)`



۷. آزمون موران

```
install.packages("ape")
```

```
library(ape)
```

```
##### محاسبه ی وزن ها و سپس آزمون موران #####
```

```
data.dists <- as.matrix(dist(cbind(data$x , data$y)))
```

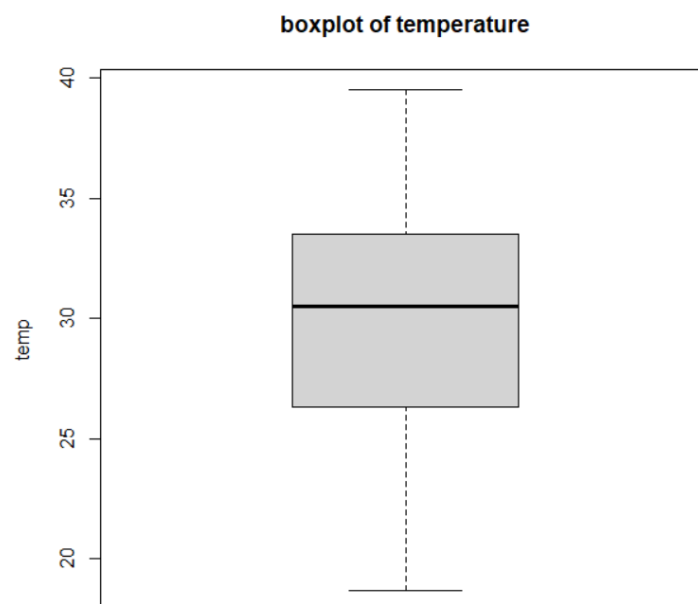
```
data.dists.bin <- (data.dists > 0 & data.dists <= .75)
```

```
Moran.I(data$z, data.dists.bin)
```

```
> Moran.I(data$z,  
$observed  
[1] 0.5150879  
  
$expected  
[1] -0.01315789  
  
$sd  
[1] 0.25947  
  
$p.value  
[1] 0.04176395
```

۸. رسم نمودار جعبه ای

```
boxplot(temp, data = data, ylab = "temp", main = "boxplot of temperature")
```



۹. اجرای رگرسیون خطی چندگانه

```
data.lm <- lm(temp ~ lon + lat, data = data)
```

```
summary(data.lm)
```

```
resid(data.lm)
```

```
plot(resid(data.lm),lon)
```

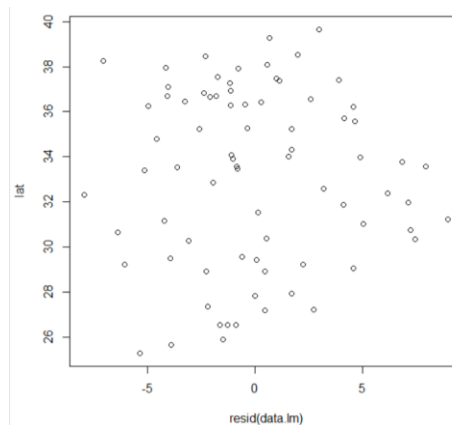
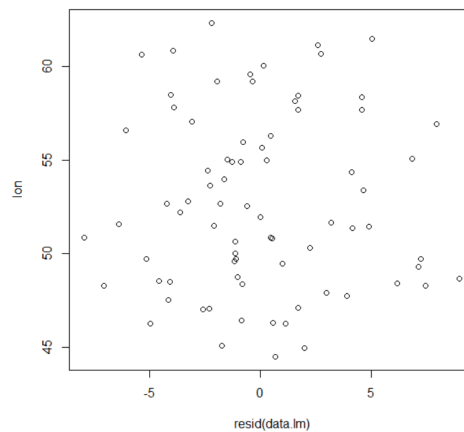
```
plot(resid(data.lm),lat)
```

```
Call:
lm(formula = temp ~ lon + lat, data = data)

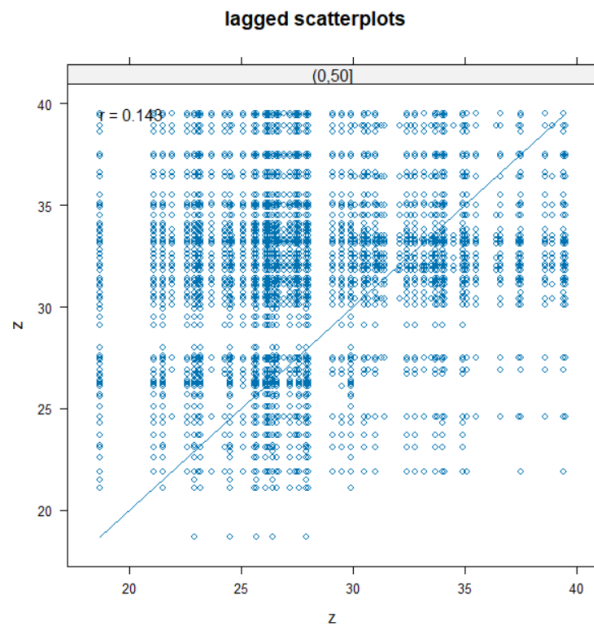
Residuals:
    Min       1Q   Median       3Q      Max
-7.9424 -2.3006 -0.6126  2.2342  8.9979

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  28.1207     8.4264   3.337 0.001326 **
lon           0.3687     0.1053   3.502 0.000787 ***
lat          -0.5270     0.1280  -4.116 9.9e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 3.842 on 74 degrees of freedom
Multiple R-squared:  0.4241,    Adjusted R-squared:  0.4085
F-statistic: 27.24 on 2 and 74 DF,  p-value: 1.361e-09
```

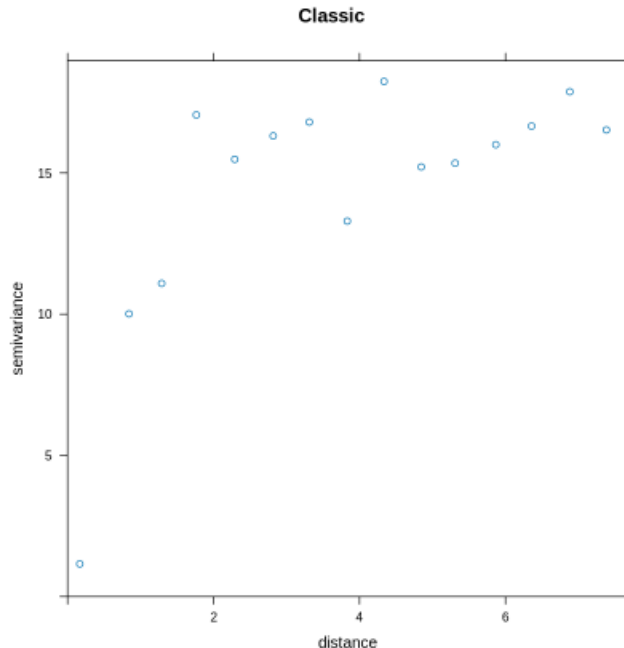


```
par(mfrow=c(2,2))  
  
install.packages("gstat")  
  
install.packages("sp")  
  
library(gstat)  
  
library(sp)  
  
coordinates(data) <- c("x", "y")  
  
hscat(z ~ 1, data=data, breaks=c(0,50,100,150,1000))
```



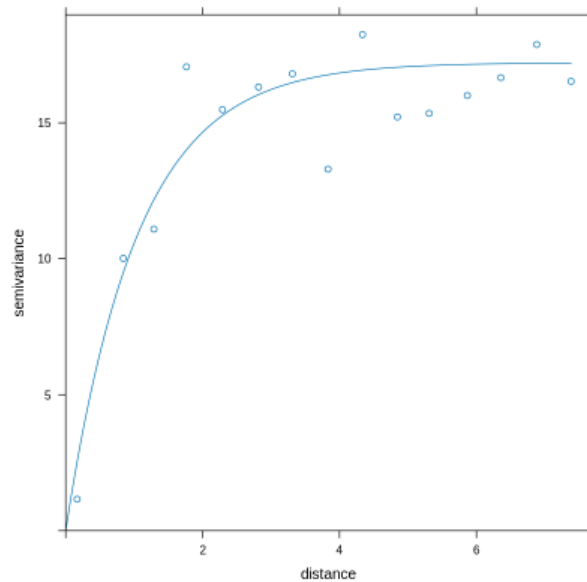


```
library(sp)
library(geoR)
library(gstat)
# تبدیل داده‌ها به دیتا فریم
latitude <- c(data_trend_removed$latitude)
longitude <- c(data_trend_removed$longitude)
Trend_Removed <- c(data_trend_removed$Trend_Removed)
data <- data.frame(latitude, longitude, Trend_Removed)
# تعریف مختصات مکانی برای دیتا فریم
coordinates(data) <- ~longitude + latitude
# محاسبه واریوگرام
var.data <- variogram(Trend_Removed ~ longitude + latitude, data = data)
# نمایش واریوگرام
plot(var.data, lty = 1, main = "Classic")
```



۱۲. انتخاب مدل تغییرنگار نمایی و برازش آن بروی داده ها:

```
library(gstat)
var1 <- variogram(Trend_Removed ~ longitude + latitude, data = data)
v.fit1 = fit.variogram(var1, vgm(50, "Exp", 2, 50))
plot(var1, v.fit1)
```



۱۳. کریگیدن عادی

```
#Kriging
library(gstat)
library(sp)
library(rgl)
library(geoR)
library(gstat)
# Load data
#data <- read.csv("/content/output.csv", encoding = "utf-8")
```

```

# Extract values
z_data <- data_trend_removed$Trend_Removed
# Create coordinates and values arrays
coords <- as.data.frame(data_trend_removed[, c("latitude", "longitude")])
values <- as.numeric(data_trend_removed$Trend_Removed)
# Load data
excel_file <- "/content/Data-new.xls"
# Read data
newdata <- read_excel(excel_file)
# Calculate distance matrix
dist_matrix <- as.matrix(dist(coords, method = "euclidean"))
max_dist <- max(dist_matrix)
H=dist(data_trend_removed)
L=max(H)
va2<-variog(G2,max.dist=L,estimator.type="classic")
m2 <- vgm(15.18, "Exp", 45.18, 1.9)
or.kriling.2 <- krige(id="z_data", formula=z_data~1, data=data_trend_removed,
                    newdata=newdata[,2:3], model = m2 ,locations=~latitude+longitude)
Un.kriling.2 <- krige(id="z.co", formula=z.co~x+y, data=a2,
                    newdata=grd2, model = m2,locations=~x+y)
or.kriling.2

```

	latitude	longitude	z_data.pred	z_data.var
	<dbl>	<dbl>	<dbl>	<dbl>
1	29.23805	50.31931	1.7653992	2.391520
2	34.20406	47.75966	-0.4824001	2.389223
3	32.77917	58.88586	0.8450770	2.515840
4	35.69395	52.04722	0.2803187	2.414864

## منابع

(۱) [آمار فضایی-د.محسن محمدزاده]

[https://pub.modares.ac.ir/book\\_treasure.php?mod=viewbook&book\\_id=29&slc\\_lang=fa&sid=1](https://pub.modares.ac.ir/book_treasure.php?mod=viewbook&book_id=29&slc_lang=fa&sid=1)

(۲) [منبع داده]

<http://www.meteomanz.com/index?l=1>

(۳) [تاریخچه آمار فضایی]

[https://fa.wikipedia.org/wiki/%D8%AA%D8%AD%D9%84%DB%8C%D9%84\\_%D9%81%D8%B6%D8%A7%DB%8C%DB%8C](https://fa.wikipedia.org/wiki/%D8%AA%D8%AD%D9%84%DB%8C%D9%84_%D9%81%D8%B6%D8%A7%DB%8C%DB%8C)