作业一

总价最贵：

	名称	地理位置 1	地理位置 2	地理位置 3	房型	面积	均价	总价
56	北京壹号总部	大兴	亦庄	台湖镇光机电一体化产业基地科创东二街 5 号	1 室	3127	28000	8755.6

总价最便宜：

	名称	地理位置 1	地理位置 2	地理位置 3	房型	面积	均价	总价
160	长海御墅	房山	房山其它	长沟国家湿地公园南侧	1 室	70	15000	105.0

总价中位数: 558.75

均价最贵:

	名称	地理位置 1	地理位置 2	地理位置 3	房型	面积	均价	总价
57	北京庄园	顺义	顺义其它	京承高速第 11 出口往东 800 米	4 室	460	167000	7682.0

均价最便宜:

	名称	地理位置 1	地理位置 2	地理位置 3	房型	面积	均价	总价
160	长海御墅	房山	房山其它	长沟国家湿地公园南侧	1 室	70	15000	105.0

均价中位数: 46000.0

总价在均值三倍标准差以外的房屋:

	名称	地理位置 1	地理位置 2	地理位置 3	房型	面积	均价	总价
57	北京庄园	顺义	顺义其它	京承高速第 11 出口往东 800 米	4 室	460	167000	7682.0
90	懋源·璟岳	丰台	玉泉营	南三环西路 99 号院	4 室	465	140000	6510.0
111	润泽御府	朝阳	北苑	北京市朝阳区北五环顾家庄桥向北约 2.6 公里	4 室	540	110000	5940.0
79	天润福熙大道	朝阳	北苑	清河营东路 1 号院, 清河营东路 3 号院	6 室	436	110000	4796.0

原因: 贫富差距大，为少部分先富起来的有钱人所建的房屋价格远高于人民群众的普遍生活水平

均价为异常值的房屋:

	名称	地理位置 1	地理位置 2	地理位置 3	房型	面积	均价	总价
57	北京庄园	顺义	顺义其它	京承高速第 11 出口往东 800 米	4 室	460	167000	7682.0
90	懋源·璟岳	丰台	玉泉营	南三环西路 99 号院	4 室	465	140000	6510.0
86	尊悦光华	朝阳	CBD	北京市朝阳区光华东里甲 1 号院 3 号楼	3 室	133	130000	1729.0
125	紫辰院	丰台	岳各庄	岳各庄北桥东北角 200 米处	5 室	266	128000	3404.8

原因: 贫富差距大，为少部分先富起来的有钱人所建的房屋价格远高于人民群众的普遍生活水平

离散化:

(0.0, 100000.0]	162	93%
(100000.0, inf]	12	7%

原因: 按均价是否大于十万设置

CSV：

名称,地理位置1,地理位置2,地理位置3,房型,面积,均价,总价

K2十里春风,通州,通州其它,永乐店镇潞小路百菜玛工业园对面,,155,28000,434.0

K2十里春风,通州,通州其它,永乐店镇潞小路百菜玛工业园对面,2室,74,24500,181.3

万年广阳郡九号,房山,长阳,长阳清苑南街与汇商东路交汇处西北角,3室,166,50000,830.0

万科弗农小镇,密云,溪翁庄镇,密关路西侧（密云水库南岸2公里）,3室,140,25000,350.0

万科雲庐,丰台,丰台其它,魏各庄路万科雲庐,4室,135,37000,499.5

万科雲庐,丰台,丰台其它,魏各庄路万科雲庐,4室,200,37000,740.0

世茂北京天誉,丰台,十里河,北京市丰台区小红门路312号,3室,145,120000,1740.0

东叁金茂府,丰台,十里河,东三环分钟寺桥南约500米,4室,125,107500,1343.75

中冶德贤公馆,大兴,旧宫,德贤东路6号院（南四环榴乡桥东南角800米）,0室,134,77000,1031.8

中国铁建·梧桐浅山,平谷,平谷其它,北京市平谷区东南路夏各庄卫生院对面,2室,90,15000,135.0

中国铁建澜廷墅,房山,房山其它,阎吕路中国铁建·山语澜廷,4室,223,28000,624.4

中国铁建澜廷墅,房山,房山其它,阎吕路南侧（北京市第十三中学西北300米处）,3室,75,32000,240.0

中国铁建花语金郡,大兴,瀛海,南海子公园西侧(南五环旧忠桥向南第二个红绿灯西300米),3室,150,70000,1050.0

中建·京西印玥,房山,良乡,阳光北大街中建·京西印玥,2室,73,43000,313.9

中建·长安麓府,门头沟,冯村,泰安路中建·长安麓府,2室,70,58000,406.0

中建国望府,丰台,丰台其它,射击场路中建国望府,5室,314,77547,2435

中海丽春湖墅·合院,昌平,沙河,地铁昌平线沙河地铁站南600米,4室,263,36000,946.8

中海云熙,大兴,大兴其它,魏善庄新城南中轴路东侧500米,2室,76,37700,286.52

中海云筑,大兴,大兴新机场,京开高速庞各庄桥西1500米，团结路北（庞各庄镇宏轩饺子馆儿对面）,3室,266,37000,984.2

中海云筑,大兴,大兴新机场,北京市大兴区团结路,3室,89,37000,329.3

中海叁號院,大兴,大兴其它,博兴八路与兴海路交叉口,3室,168,70000,1176.0

中海寰宇视界,房山,长阳,北京市房山区轨道交通房山线稻田地铁站南侧200米处,1室,53,58000,307.4

中海甲叁號院,丰台,玉泉营,丰台恒丰路,3室,145,109000,1580.5

中海首开拾光里,朝阳,中央别墅区,朝阳区金盏南路，金港国际赛车场东侧,3室,140,70000,980.0

中粮京西祥云,房山,长阳,地铁稻田站北800米，西邻京深路,4室,115,58000,667.0

中粮天恒天悦壹号,丰台,西红门,南四环地铁新宫站南500米,4室,220,85000,1870.0

中铁华侨城和园,大兴,瀛海,南五环南海子公园西侧约500米,3室,154,60000,924.0

中铁诺德阅墅,顺义,中央别墅区,顺义区后沙峪优山美地D区西北侧,4室,220,62000,1364.0

中骏云景台,房山,房山其它,官道西大街与良常路交叉口,2室,75,28000,210.0

中骏西山天璟,门头沟,城子,西山永定楼北300米,4室,117,65000,760.5

亦庄橡树湾,通州,马驹桥,马驹桥镇亦庄新城潞马路与兴华南街交汇处东300米,1室,66,45000,297.0

京投发展·锦悦府,密云,鼓楼街道,檀营乡檀东路西侧,3室,187,25607,478.8509

京投发展·锦悦府,密云,鼓楼街道,檀营乡檀东路西侧,3室,79,25607,202.2953

京西悦府,房山,阎村,北京市房山区燕房线阎村地铁站东南角约189米,3室,175,50000,875.0

京西悦府,房山,阎村,燕房线阎村地铁站东南角约189米,3室,120,33000,396.0

京贸国际公馆,通州,九棵树(家乐福),怡乐中路299号院（广渠快速路二期出口向南1000米）,1室,72,64000,460.8

京贸国际城·峰景,通州,武夷花园,芙蓉东路1号（通燕高速耿庄桥北出口向南300米）,1室,69,68000,469.2

住总兴创如遇,大兴,大兴新机场洋房别墅区,采华路与育镇街交叉路口往西北约100米(采育文化广场西南侧约100米),1室,53,27000,143.1

住总如院,大兴,大兴新机场洋房别墅区,北京市大兴区采华路(波尔多小镇南区西南侧约250米),2室,98,31000,303.8

住总如院,大兴,大兴新机场洋房别墅区,北京市大兴区采华路(波尔多小镇南区西南侧约250米),3室,240,22728,545.472

住总正华·时代广场,大兴,天宫院,地铁4号线生物医药基地站南200米,0室,61,55000,335.5

保利绿城·和锦诚园,大兴,瀛海,8 号线南段地铁站东 1.6 公里,4 室,130,63000,819.0

保利锦上,朝阳,垡头,化工路保利锦上,3 室,114,71000,809.4

元熙华府,丰台,宋家庄,东南三环东铁营桥向南 600 米,3 室,126,87000,1096.2

公园十七区,顺义,后沙峪,顺义区后沙峪火沙路和裕庆路交汇口北 500 米,4 室,90,55000,495.0

公园都会,顺义,顺义其它,北京顺义区顺泰路与外环路交汇处,2 室,72,39000,280.8

兴创荣墅,大兴,大兴新机场洋房别墅区,北京市大兴区育胜街,3 室,240,23000,552.0

凯德麓语,昌平,昌平其它,兴寿镇京承高速 G11 出口向西怀昌路北侧,3 室,280,35000,980.0

北京东湾,通州,万达,通惠北路 98 号,1 室,58,69600,403.68

作业二

存在的空值列：

column PM_Dongsi has 164 null values

column PM_Dongsihuan has 3295 null values

column PM_Nongzhanguan has 287 null values

column PM_US Post has 129 null values

column DEWP has 5 null values

column HUMI has 339 null values

column PRES has 339 null values

column TEMP has 5 null values

column cbwd has 5 null values

column lws has 5 null values

column precipitation has 459 null values

column Iprec has 459 null values

处理方法： 因为数据的变化是连续有规律的，使用线性插值较为准确，对于无法进行插值的风向信息设置为 unknown

CSV：

,No,year,month,day,hour,season,PM_Dongsi,PM_Dongsihuan,PM_Nongzhanguan,PM_US Post,DEWP,HUMI,PRES,TEMP,cbwd,lws,precipitation,Iprec
43824,43825,2015,1,1,0,4,5.0,32.0,8.0,22.0,-21.0,29.0,1034.0,-6.0,SE,0.89,0.0,0.0
43825,43826,2015,1,1,1,4,4.0,12.0,7.0,9.0,-22.0,23.0,1034.0,-4.0,NW,4.92,0.0,0.0
43826,43827,2015,1,1,2,4,3.0,19.0,7.0,9.0,-21.0,27.0,1034.0,-5.0,NW,8.94,0.0,0.0
43827,43828,2015,1,1,3,4,4.0,9.0,11.0,13.0,-21.0,29.0,1035.0,-6.0,NW,12.96,0.0,0.0
43828,43829,2015,1,1,4,4,3.0,11.0,5.0,10.0,-21.0,27.0,1034.0,-5.0,NW,16.98,0.0,0.0
43829,43830,2015,1,1,5,4,3.0,18.0,3.0,6.0,-22.0,23.0,1034.0,-4.0,NW,24.13,0.0,0.0
43830,43831,2015,1,1,6,4,3.0,20.0,6.0,8.0,-23.0,22.0,1034.0,-5.0,NW,25.92,0.0,0.0
43831,43832,2015,1,1,7,4,3.0,22.0,7.0,17.0,-22.0,26.0,1035.0,-6.0,SE,1.79,0.0,0.0
43832,43833,2015,1,1,8,4,4.0,29.5,9.0,11.0,-22.0,29.0,1035.0,-7.0,cv,0.89,0.0,0.0
43833,43834,2015,1,1,9,4,5.0,37.0,11.0,33.0,-22.0,24.0,1035.0,-5.0,NE,1.79,0.0,0.0
43834,43835,2015,1,1,10,4,4.0,37.0,36.0,37.0,-22.0,21.0,1035.0,-3.0,NE,4.92,0.0,0.0
43835,43836,2015,1,1,11,4,21.0,40.0,40.0,40.0,-22.0,19.0,1034.0,-2.0,cv,1.79,0.0,0.0
43836,43837,2015,1,1,12,4,41.0,63.0,61.0,63.0,-22.0,17.0,1032.0,0.0,cv,3.58,0.0,0.0

43837,43838,2015,1,1,13,4,40.0,58.0,54.0,62.0,-22.0,16.0,1030.0,1.0,SE,3.13,0.0,0.0
43838,43839,2015,1,1,14,4,28.0,48.0,53.0,44.0,-23.0,13.0,1029.0,2.0,SE,6.26,0.0,0.0
43839,43840,2015,1,1,15,4,29.0,42.0,41.0,48.0,-23.0,13.0,1028.0,2.0,SE,9.39,0.0,0.0
43840,43841,2015,1,1,16,4,31.0,53.0,51.0,51.0,-24.0,12.0,1027.0,2.0,SE,13.41,0.0,0.0
43841,43842,2015,1,1,17,4,52.0,68.0,68.0,82.0,-23.0,14.0,1027.0,1.0,SE,16.54,0.0,0.0
43842,43843,2015,1,1,18,4,64.0,85.0,81.0,87.0,-21.0,20.0,1026.0,-1.0,SE,19.67,0.0,0.0
43843,43844,2015,1,1,19,4,75.0,94.0,88.0,106.0,-19.0,25.0,1026.0,-2.0,cv,0.89,0.0,0.0
43844,43845,2015,1,1,20,4,82.0,107.0,100.0,123.0,-19.0,34.0,1026.0,-6.0,NE,1.79,0.0,0.0
43845,43846,2015,1,1,21,4,88.0,138.0,102.0,136.0,-19.0,40.0,1026.0,-8.0,NE,2.68,0.0,0.0
43846,43847,2015,1,1,22,4,86.0,158.0,124.0,139.0,-18.0,38.0,1026.0,-6.0,NW,1.79,0.0,0.0
43847,43848,2015,1,1,23,4,80.0,175.0,134.0,154.0,-17.0,48.0,1027.0,-8.0,NE,1.79,0.0,0.0
43848,43849,2015,1,2,0,4,82.0,161.0,126.0,126.0,-18.0,32.0,1027.0,-4.0,NW,1.79,0.0,0.0
43849,43850,2015,1,2,1,4,81.0,119.0,98.0,98.0,-19.0,32.0,1028.0,-5.0,NW,4.92,0.0,0.0
43850,43851,2015,1,2,2,4,68.0,95.0,68.0,66.0,-18.0,35.0,1028.0,-5.0,NW,9.84,0.0,0.0
43851,43852,2015,1,2,3,4,35.0,52.0,47.0,45.0,-18.0,28.0,1029.0,-2.0,NE,4.92,0.0,0.0
43852,43853,2015,1,2,4,4,16.0,27.0,27.0,28.0,-18.0,30.0,1030.0,-3.0,NE,8.94,0.0,0.0
43853,43854,2015,1,2,5,4,8.0,18.0,12.0,12.0,-18.0,30.0,1030.0,-3.0,NE,12.07,0.0,0.0
43854,43855,2015,1,2,6,4,5.0,20.0,13.0,12.0,-18.0,35.0,1031.0,-5.0,cv,0.89,0.0,0.0
43855,43856,2015,1,2,7,4,3.0,20.0,12.0,16.0,-17.0,44.0,1031.0,-7.0,NE,1.79,0.0,0.0
43856,43857,2015,1,2,8,4,3.0,25.0,12.0,13.0,-17.0,41.0,1032.0,-6.0,NW,3.13,0.0,0.0
43857,43858,2015,1,2,9,4,8.0,27.0,18.0,24.0,-15.0,33.0,1033.0,-1.0,cv,0.89,0.0,0.0
43858,43859,2015,1,2,10,4,11.0,29.0,21.0,34.0,-18.0,24.0,1033.0,0.0,NE,3.13,0.0,0.0
43859,43860,2015,1,2,11,4,8.0,18.0,14.0,13.0,-18.0,21.0,1032.0,2.0,NE,6.26,0.0,0.0
43860,43861,2015,1,2,12,4,9.0,22.0,11.0,19.0,-19.0,18.0,1031.0,3.0,NE,10.28,0.0,0.0
43861,43862,2015,1,2,13,4,9.0,25.0,16.0,26.0,-18.0,17.0,1030.0,5.0,NE,12.07,0.0,0.0
43862,43863,2015,1,2,14,4,12.0,21.0,21.0,26.0,-19.0,15.0,1029.0,5.0,SE,1.79,0.0,0.0
43863,43864,2015,1,2,15,4,28.0,34.0,31.0,41.0,-19.0,15.0,1029.0,5.0,SE,4.92,0.0,0.0
43864,43865,2015,1,2,16,4,55.0,63.0,59.0,85.0,-18.0,18.0,1028.0,4.0,SE,8.05,0.0,0.0
43865,43866,2015,1,2,17,4,58.666666666666664,101.0,104.0,108.0,-18.0,19.0,1028.0,3.0,SE,12.07,0.0,0.0
43866,43867,2015,1,2,18,4,62.333333333333336,89.0,86.0,87.0,-17.0,21.0,1028.0,3.0,cv,4.02,0.0,0.0
43867,43868,2015,1,2,19,4,66.0,87.0,82.0,102.0,-18.0,22.0,1028.0,1.0,SE,1.79,0.0,0.0
43868,43869,2015,1,2,20,4,115.0,109.0,103.0,117.0,-17.0,24.0,1027.0,1.0,SE,4.92,0.0,0.0
43869,43870,2015,1,2,21,4,131.0,118.0,112.0,125.0,-16.0,33.0,1027.0,-2.0,SE,6.71,0.0,0.0
43870,43871,2015,1,2,22,4,143.0,131.0,121.0,145.0,-15.0,39.0,1026.0,-3.0,SE,7.6,0.0,0.0
43871,43872,2015,1,2,23,4,159.0,150.0,132.0,157.0,-15.0,42.0,1025.0,-4.0,cv,0.89,0.0,0.0
43872,43873,2015,1,3,0,4,170.0,171.0,144.0,163.0,-15.0,45.0,1024.0,-5.0,cv,1.34,0.0,0.0