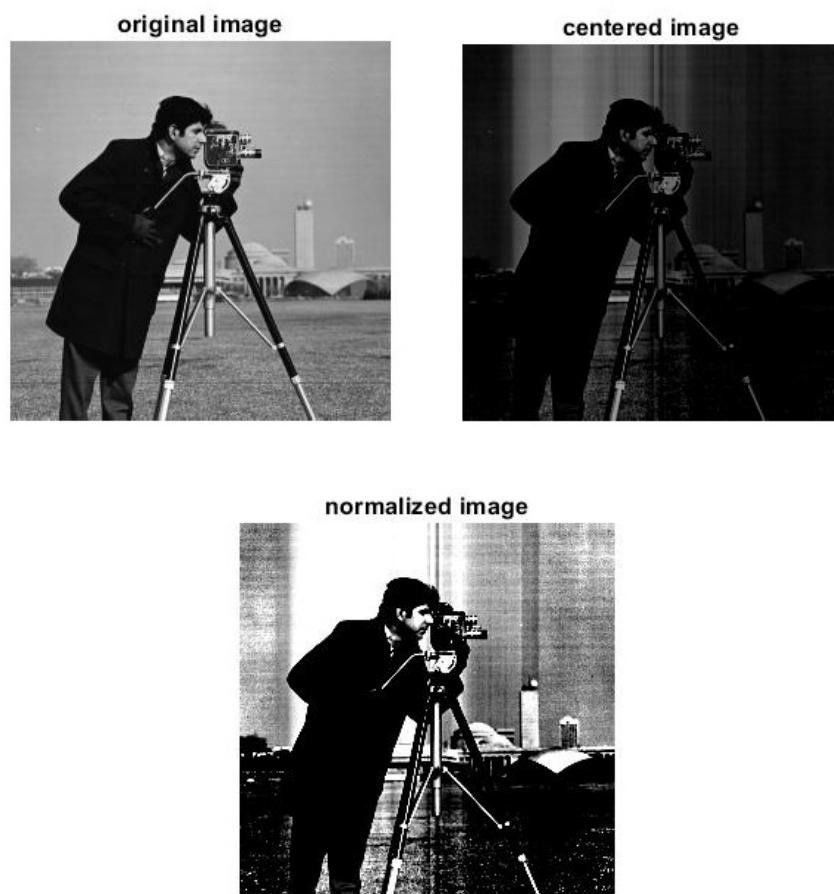
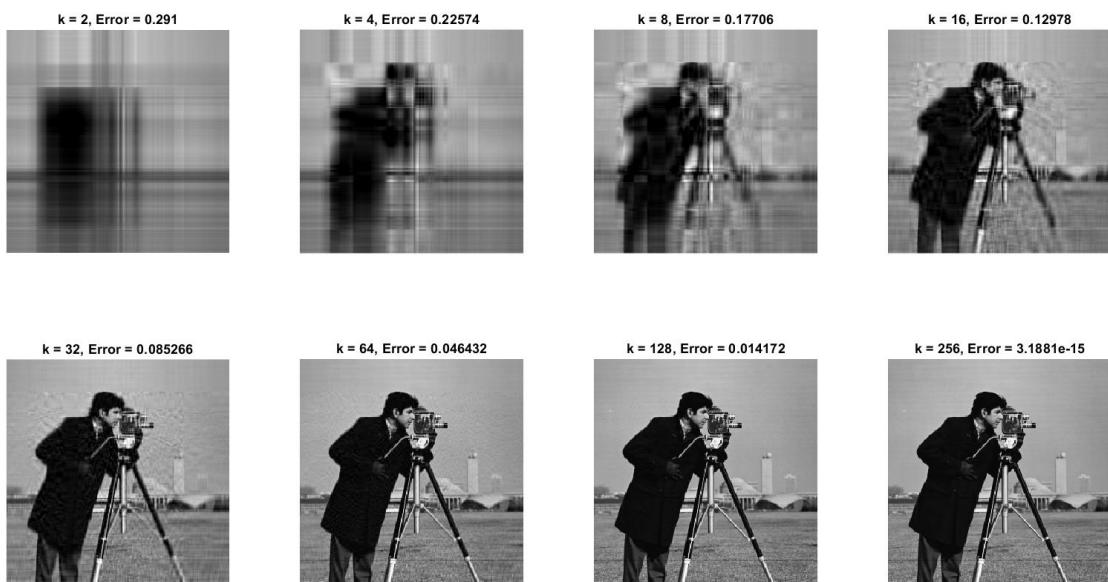


سوال ۱: ابتدا تصویر مورد نظر را می‌خوانیم و سپس دو حالت ممکن برای نرمالیزاسیون آن را نمایش می‌دهیم.

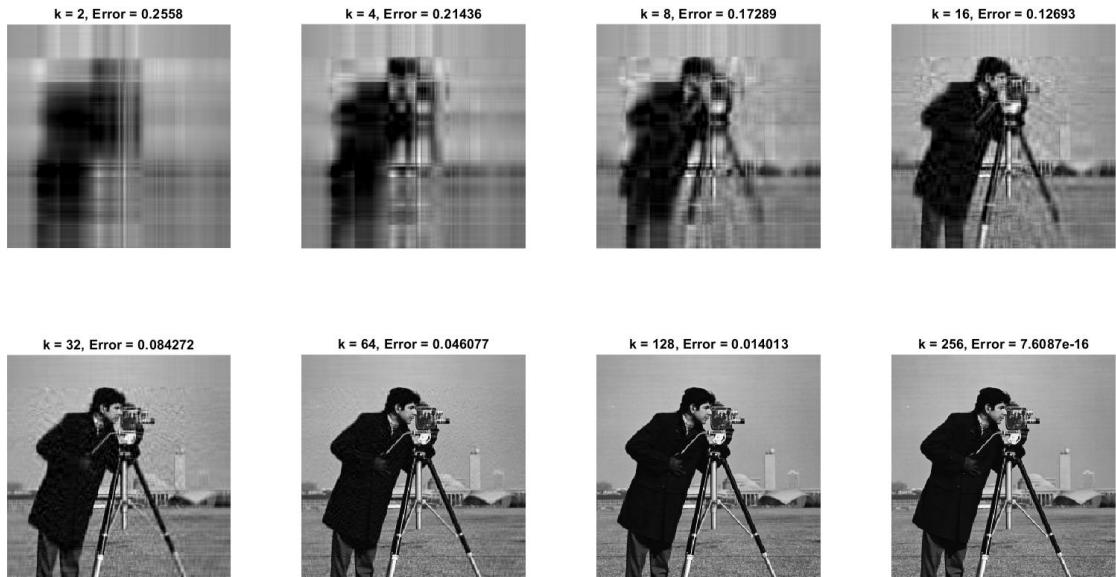


الف) برای رتبه‌های مختلف $\{1, \dots, 8, \dots, 2^k\}$ برای تجزیه truncated SVD، تصویر بازسازی شده را رسم می‌کنیم و خطای هر یک را بدست می‌آوریم.

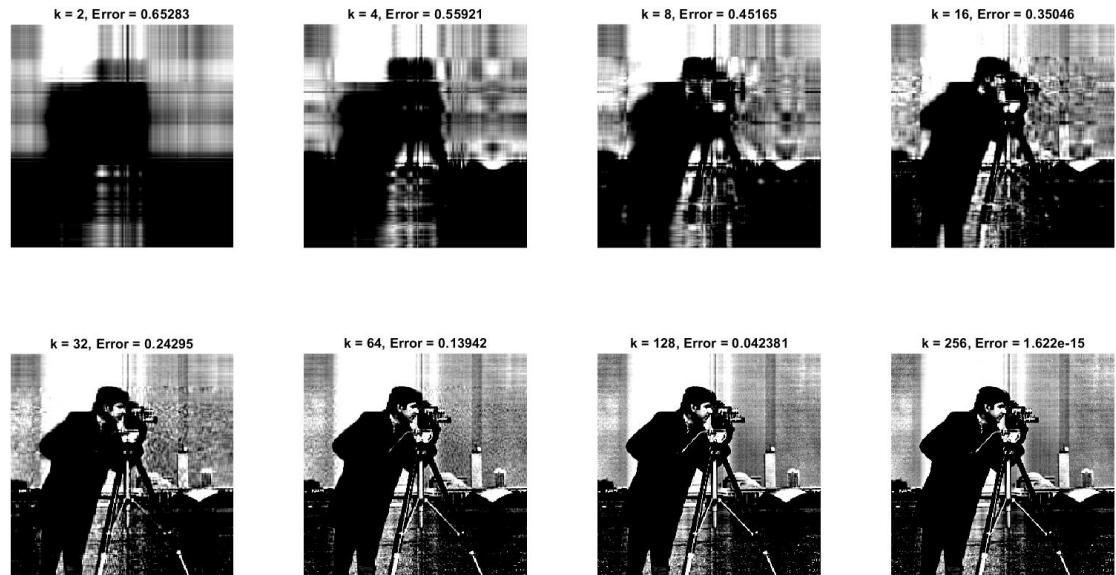
Original Image:



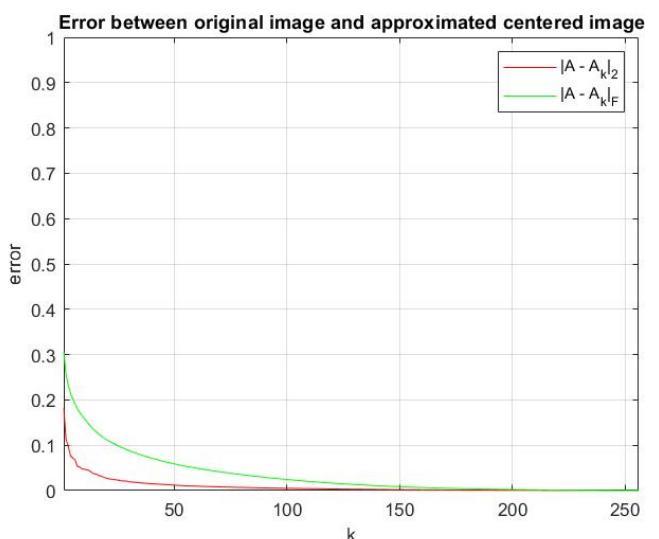
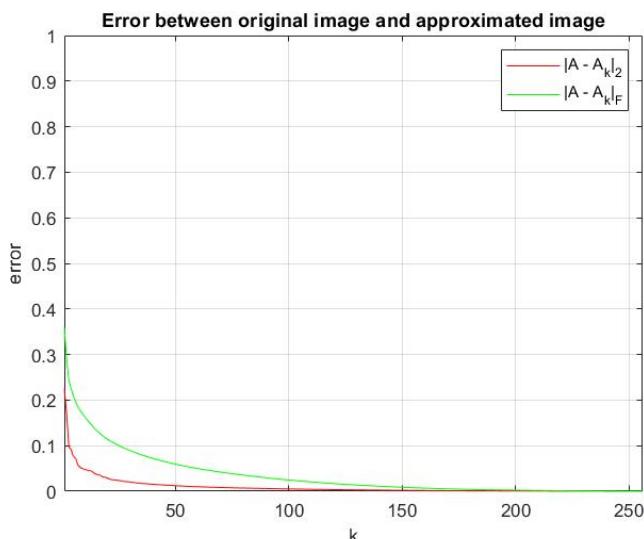
Centered Image:

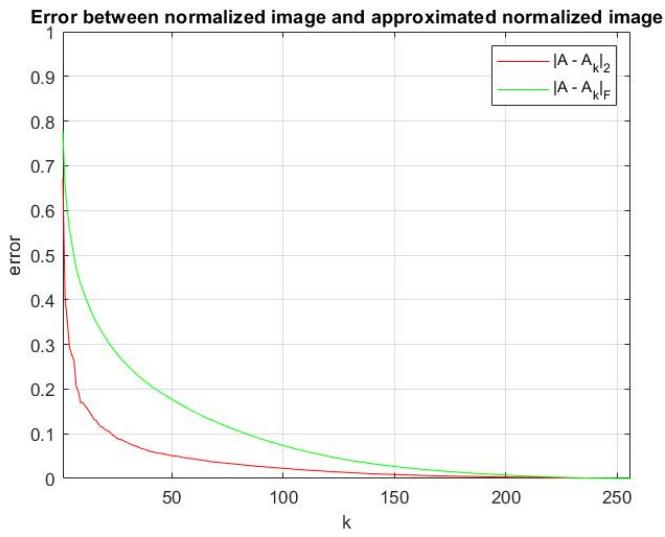


Normalized Image:

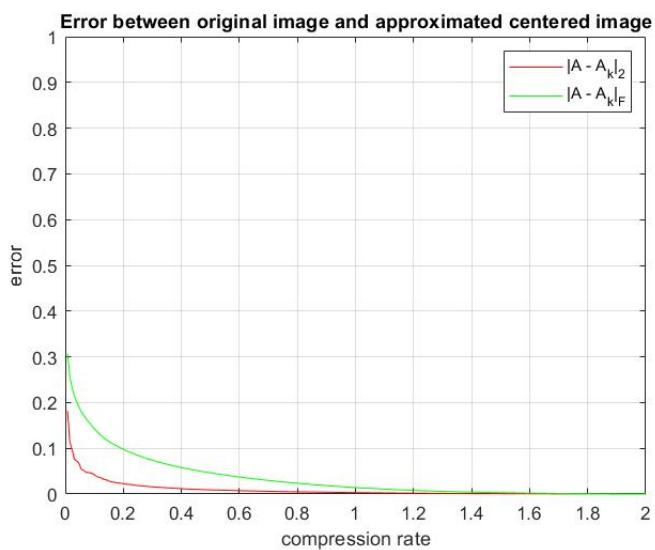
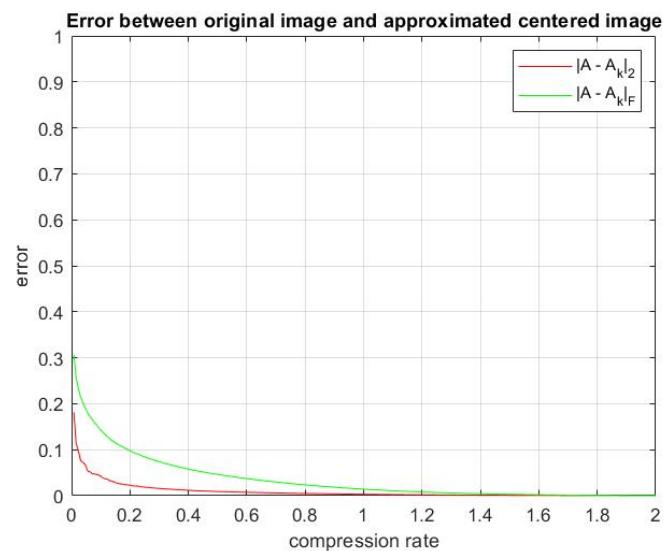
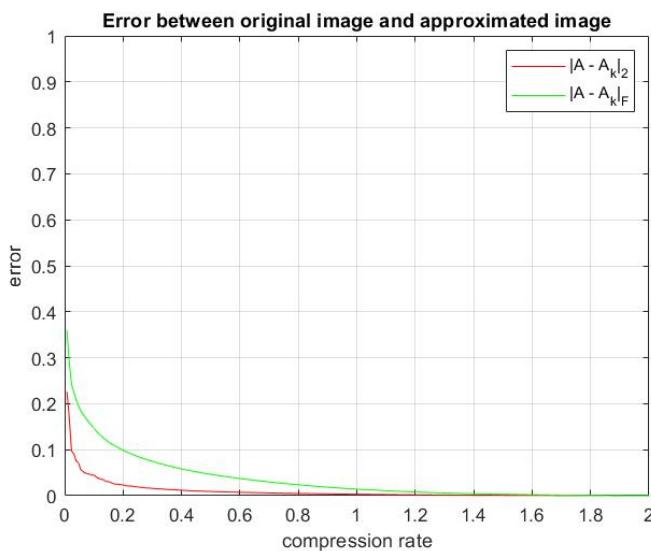


ب) نمودار خطای حسب رتبه به صورت زیر بدست می آید:





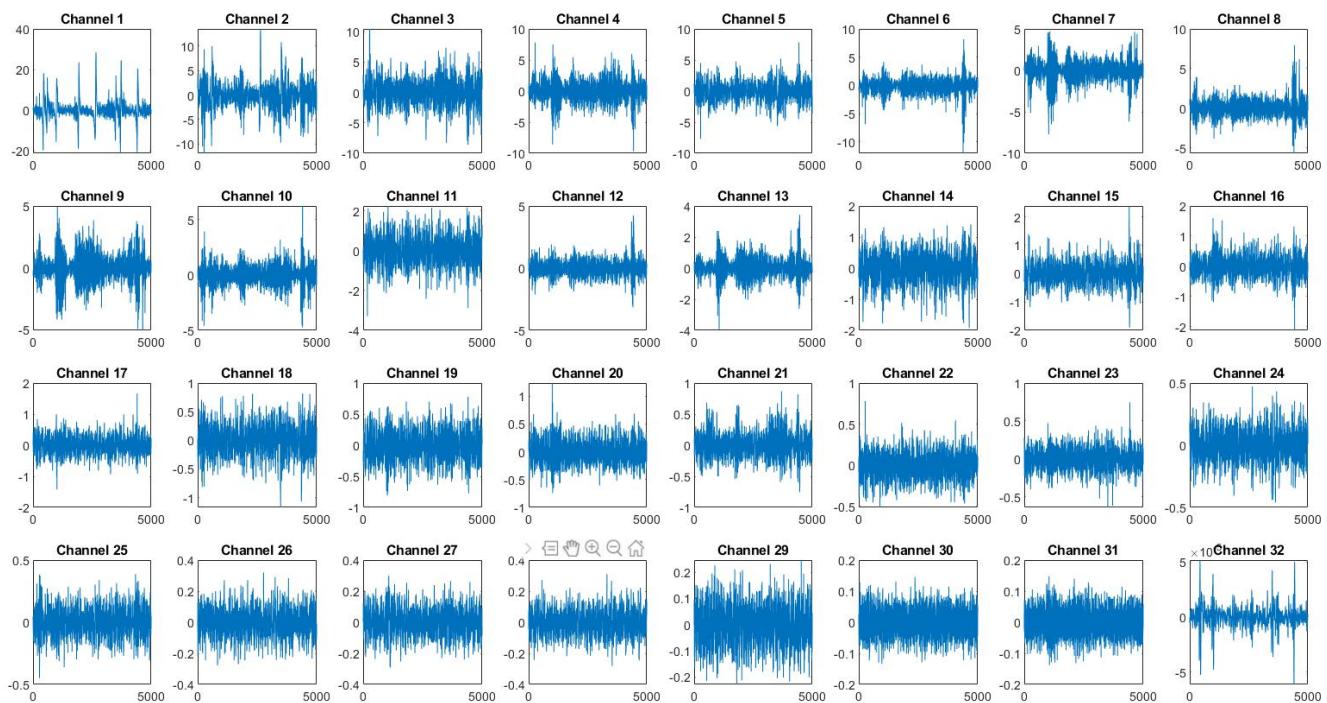
ب) نمودار خطای حسب مقدار فشرده شدن تصویر به صورت زیر بدست می‌آید:



از نمودارهای بالا نتیجه می‌شود که حذف میانگین قبل از محاسبه SVD و اضافه کردن جداگانه آن هنگام بازسازی تصویر، اندکی خطای کاهش می‌دهد.

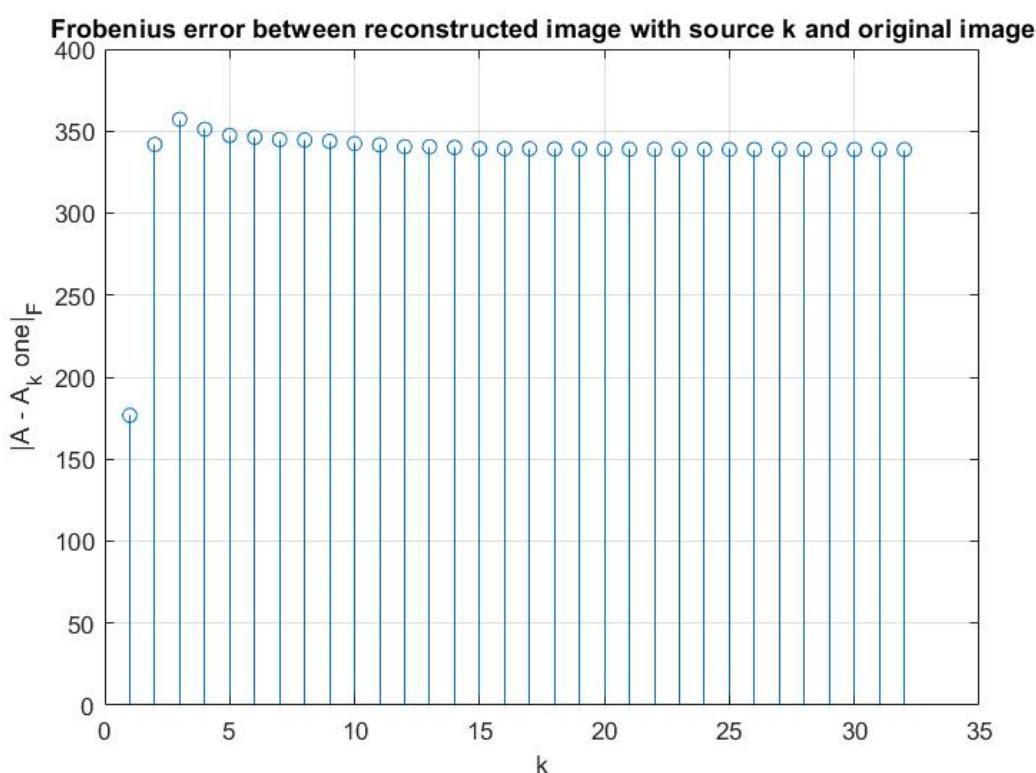
سوال ۲: ابتدا میانگین هر کanal سیگنال نویزی را از آن کم می‌کنیم و سپس SVD حاصل را بدست می‌آوریم. برای انتخاب منابع غیر نویزی سه روش را بررسی می‌کنیم.

در روش اول منابع بدست آمده را به صورت جدا گانه رسم می‌کنیم و به صورت چشمی منابع غیر نویزی را انتخاب می‌کنیم.



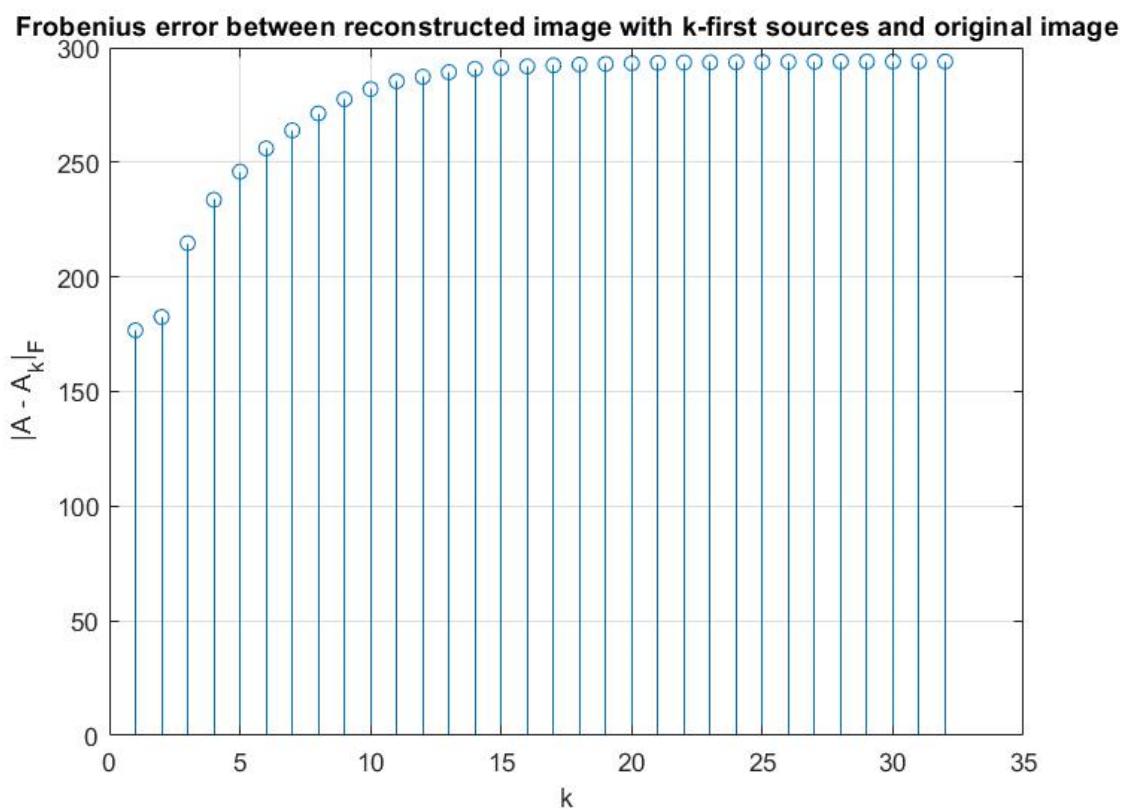
به صورت چشمی تنها منبع اول غیر نویزی است و در بقیه منابع شدت نویز بیش از حد قابل قبول است.

در روش دوم بازسازی سیگنال تنها توسط یک منبع را مد نظر قرار می‌دهیم و خطا را بدست می‌آوریم.



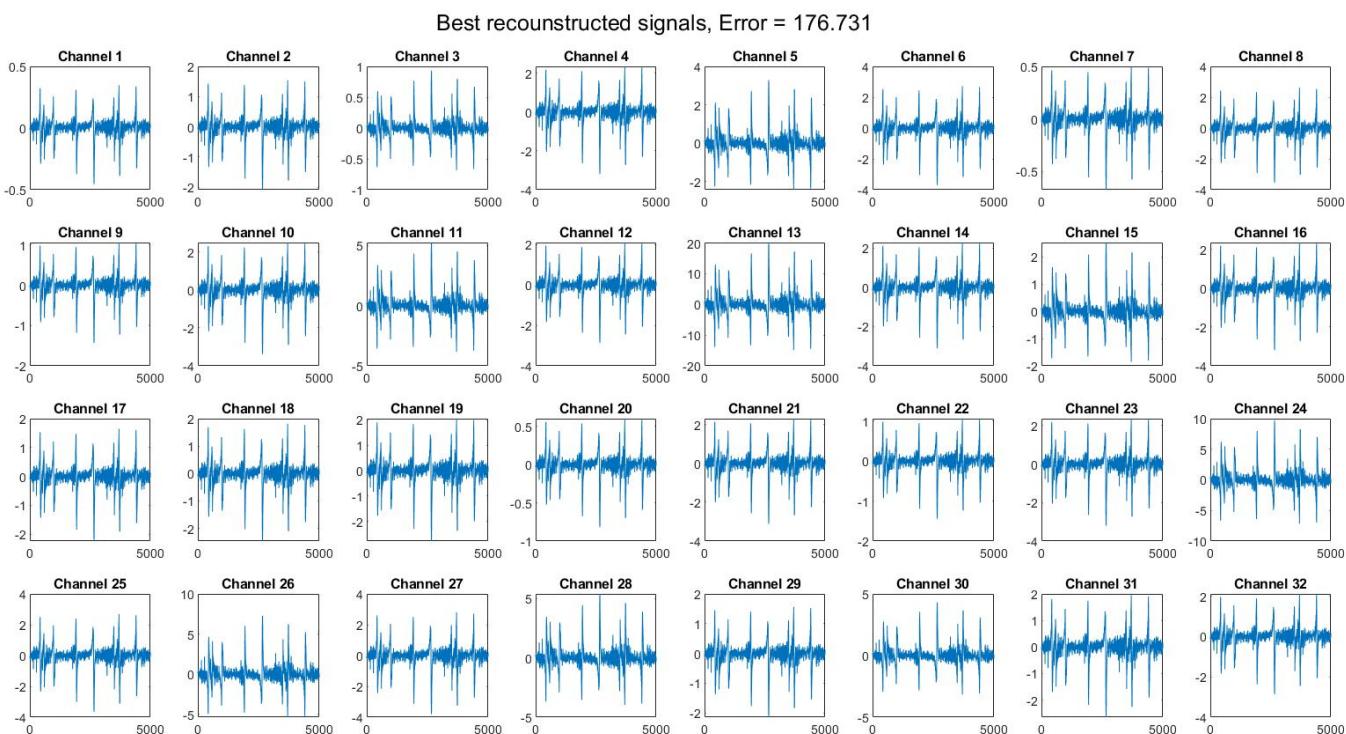
به وضوح منبع اول بهترین عملکرد را دارد که با مشاهدات روش قبل تطابق دارد.

در روش سوم بازسازی سیگنال توسط k منبع اصلی (k منبعی که بیشترین مقدار تکین را دارند) را مد نظر قرار می‌دهیم و خطای بدست می‌آوریم.

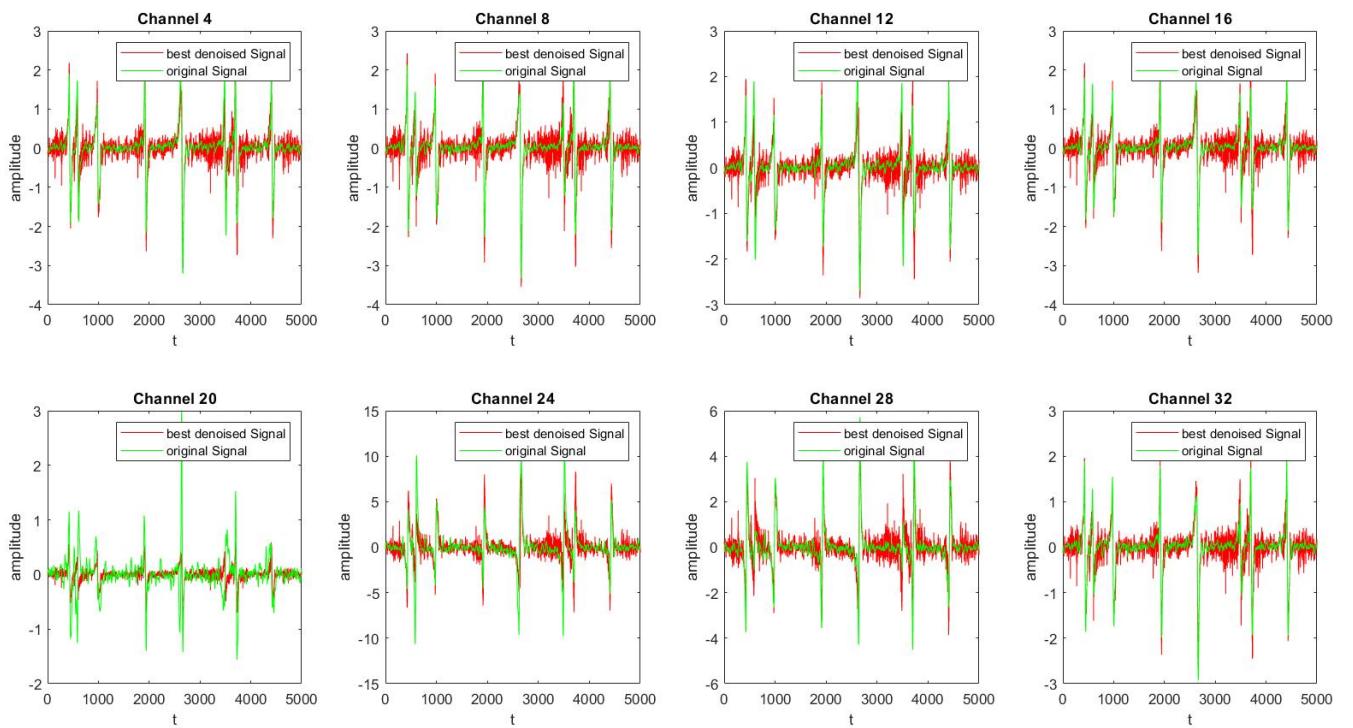


مشخصا استفاده از منابع بیشتر، نویز بازسازی را افزایش داده است. لذا با توجه به بررسی‌های انجام شده، کمترین خطای زمانی بدست می‌آید که بازسازی تنها با استفاده از منبع اول صورت گیرد.

سیگنال‌های بازسازی شده از این طریق به صورت زیر خواهند بود:

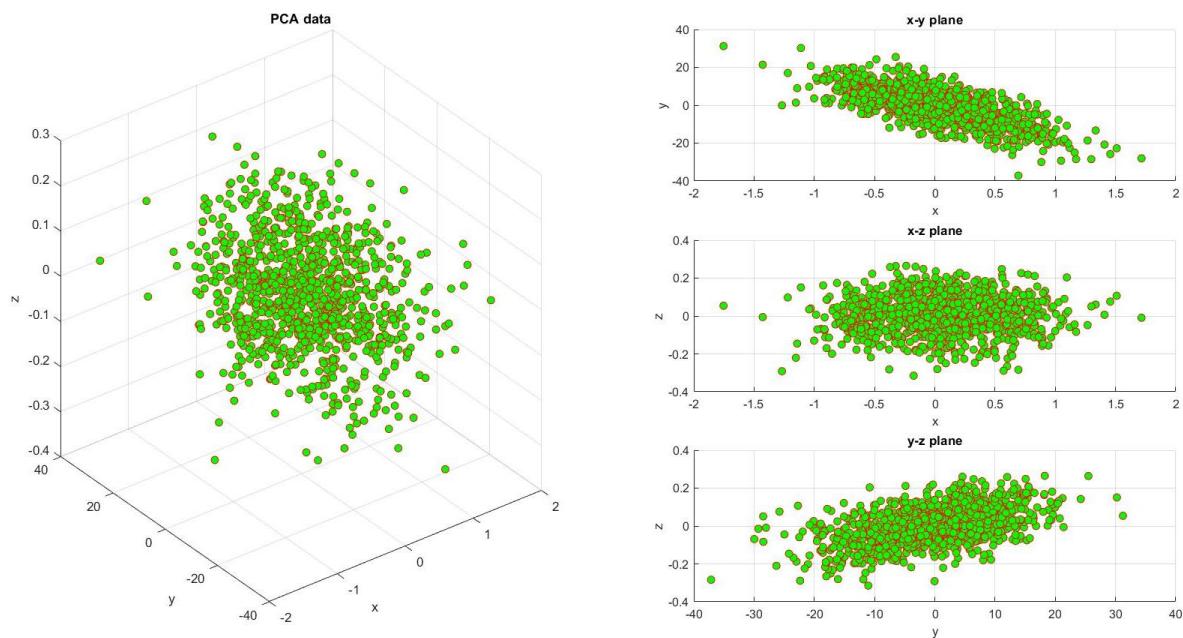


به عنوان نمونه چند سیگنال اصلی را به همراه سیگنال های بازسازی شده مرتبط نماش می دهیم.



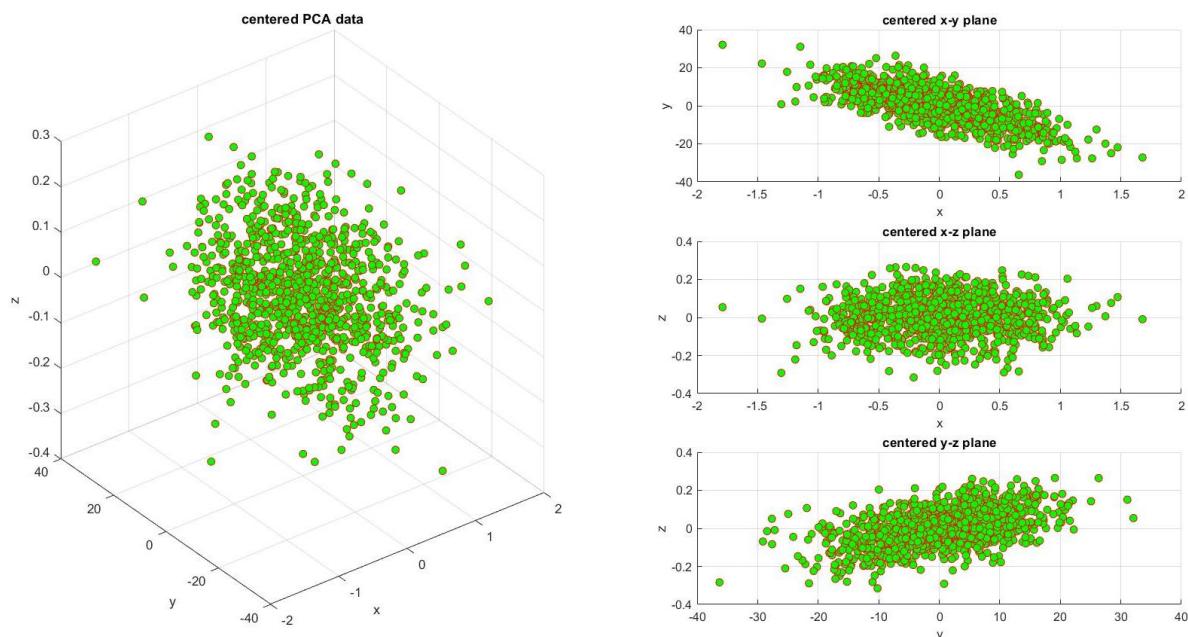
سوال ۳: الف) ابتدا مشاهدات سه بعدی مورد نظر را می خوانیم و سپس آن را به صورت های سه بعدی و دو بعدی (در صفحه های مختصات)، در دو حالت، یکی بدون حذف میانگین و دیگری با حذف آن، نمایش می دهیم.

Original data:



$$x \text{ tautness} = 15.1745, \quad y \text{ tautness} = 304.8869, \quad z \text{ tautness} = 3.0939$$

Centered data:



$$\text{centered } x \text{ tautness} = 15.1745, \quad \text{centered } y \text{ tautness} = 304.8869, \\ \text{centered } z \text{ tautness} = 3.0939$$

به وضوح کشیدگی در راستای محور y بسیار بیشتر از کشیدگی در راستای محورهای X و Z و همچنین کشیدگی در راستای محور X بسیار بیشتر از کشیدگی در راستای محور Z است.

ب) ابتدا میانگین هر محور را از آن کم می‌کنیم و سپس SVD حاصل را بدست می‌آوریم. مفاهیم مرتبط با تجزیه PCA به صورت زیر قابل محاسبه خواهد بود.

کشیدگی در راستاهای جدید: مقدار تکین متناظر با آن راستا

```
new_tautness_SVD =  
  
303.9310      0      0  
0    10.4305      0  
0      0    2.0500
```

واریانس در راستاهای جدید: مقدار تکین متناظر با آن راستا به توان تقسیم بر تعداد مشاهدات

```
new_variance_SVD =  
  
92.3740      0      0  
0    0.1088      0  
0      0    0.0042
```

جهت‌های جدید: ستون‌های ماتریس V

```
new_direction_SVD =  
  
-0.0366   -0.9842    0.1732  
0.9993   -0.0351    0.0112  
0.0049   -0.1734   -0.9848
```

داده‌های سفیدسازی شده: ستون‌های ماتریس U * S

ج) موارد قسمت قبل را با دستور PCA متلب تکرار می‌کنیم. مفاهیم مرتبط با تجزیه PCA به صورت زیر قابل محاسبه خواهد بود.

کشیدگی در راستاهای جدید: جذر حاصل ضرب متغیر بردار latent متناظر با آن راستا در تعداد مشاهدات

```
new_tautness_PCA =  
  
304.0830   10.4357    2.0510
```

واریانس در راستاهای جدید: متغیر بردار latent متناظر با آن راستا

```
new_variance_PCA =  
  
92.4665    0.1089    0.0042
```

جهت‌های جدید: ستون‌های ماتریس coeff

```
new_direction_PCA =  
  
-0.0366    0.9842   -0.1732  
0.9993    0.0351   -0.0112  
0.0049    0.1734    0.9848
```

داده‌های سفیدسازی شده: ستون‌های ماتریس score

سوال ۴: در این سوال برای بدست آوردن نتایج بهتر، ابتدا هر ویژگی مجموعه داده اصلی را به طور جداگانه نرمالیزه می‌کنیم و سپس svd حاصل را بدست می‌آوریم. در ادامه با دو دیدگاه نتیجه را بررسی می‌کنیم.

در دیدگاه اول هدف تنها بدست آوردن ویژگی‌های بهتر یا فشرده‌سازی داده است. در راستای این امر مقادیر تکین و راستاهای جدید را مورد بررسی قرار می‌دهیم.

house_S =

55.6243	0	0	0	0	0	0	0	0	0	0	0	0
0	26.9036	0	0	0	0	0	0	0	0	0	0	0
0	0	25.0504	0	0	0	0	0	0	0	0	0	0
0	0	0	20.8105	0	0	0	0	0	0	0	0	0
0	0	0	0	20.5325	0	0	0	0	0	0	0	0
0	0	0	0	0	18.2206	0	0	0	0	0	0	0
0	0	0	0	0	0	16.4425	0	0	0	0	0	0
0	0	0	0	0	0	0	14.1432	0	0	0	0	0
0	0	0	0	0	0	0	0	11.8261	0	0	0	0
0	0	0	0	0	0	0	0	0	10.5461	0	0	0
0	0	0	0	0	0	0	0	0	0	9.6921	0	0
0	0	0	0	0	0	0	0	0	0	0	9.2465	0
0	0	0	0	0	0	0	0	0	0	0	0	5.6632

house_V =

0.2510	-0.3153	0.2466	-0.0618	0.0822	-0.2197	0.7776	-0.1534	0.2604	-0.0194	-0.1096	-0.0868	0.0460
-0.2563	-0.3233	0.2959	-0.1287	0.3206	-0.3234	-0.2750	0.4027	0.3581	-0.2675	0.2628	0.0714	-0.0809
0.3467	0.1125	-0.0159	-0.0171	-0.0078	-0.0761	-0.3396	-0.1739	0.6444	0.3635	-0.3032	0.1132	-0.2511
0.0050	0.4548	0.2898	-0.8159	0.0865	0.1675	0.0741	0.0247	-0.0137	0.0062	0.0139	0.0040	0.0359
0.3429	0.2191	0.1210	0.1282	0.1369	-0.1530	-0.1996	-0.0801	-0.0185	-0.2311	0.1113	-0.8043	0.0436
-0.1892	0.1493	0.5940	0.2806	-0.4234	0.0593	0.0639	0.3268	0.0479	0.4314	0.0532	-0.1529	0.0456
0.3137	0.3120	-0.0177	0.1752	0.0167	-0.0717	0.1160	0.6008	-0.0676	-0.3628	-0.4592	0.2119	-0.0386
-0.3215	-0.3491	-0.0497	-0.2154	0.0986	0.0234	-0.1039	0.1218	-0.1533	0.1712	-0.6957	-0.3909	-0.0183
0.3198	-0.2715	0.2873	-0.1323	-0.2041	-0.1432	-0.1379	-0.0804	-0.4709	-0.0219	0.0365	0.1070	-0.6335
0.3385	-0.2395	0.2207	-0.1033	-0.1305	-0.1929	-0.3149	-0.0828	-0.1766	0.0352	-0.1048	0.2152	0.7202
0.2049	-0.3059	-0.3234	-0.2826	-0.5840	0.2732	0.0023	0.3179	0.2544	-0.1534	0.1745	-0.2096	0.0234
-0.2030	0.2386	-0.3001	-0.1685	-0.3456	-0.8035	0.0703	0.0049	-0.0449	0.0965	0.0193	-0.0417	-0.0045
0.3098	-0.0743	-0.2670	-0.0694	0.3946	-0.0532	0.0870	0.4244	-0.1952	0.6007	0.2714	-0.0552	0.0244

با توجه به این مقادیر، قسمت معناداری از تغییرات موجود در مجموعه داده اصلی در راستای ویژگی اول جدید است، چون مقدار تکین متناظر با آن اختلاف قابل توجهی با دیگر مقدارهای تکین دارد. این ویژگی در واقع برآیندی از همه ویژگی‌های مجموعه داده اصلی است و تقریباً به میزانی یکسان از همه این ویژگی‌ها تاثیر می‌پذیرد. همچنین نتیجه می‌شود همبستگی بین ویژگی‌های مختلف مجموعه داده اصلی خیلی زیاد نیست، زیرا سرعت کوچک شدن مقدارهای تکین کند است و مقدار تکین آخر تنها ۱۰٪ مقدار تکین اول است.

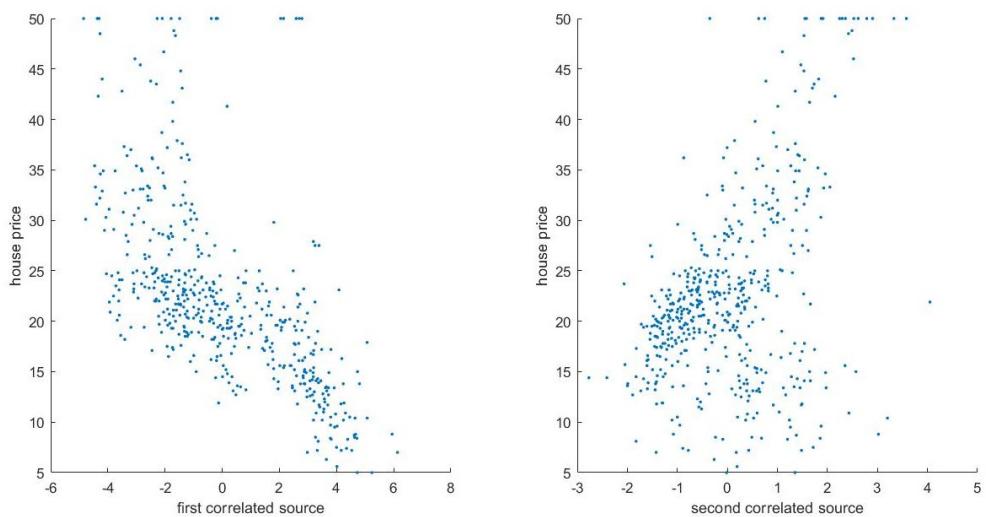
در دیدگاه دوم هدف بدست آوردن ویژگی‌های مناسب برای بدست آوردن قیمت خانه است، لذا ضریب همبستگی قیمت خانه با هر یک از ستون‌های ماتریس U را محاسبه می‌کنیم.

covariance =

-0.2310	0.1079	0.1602	0.0411	-0.0838	-0.0223	-0.0028	-0.0269	0.0032	-0.0214	0.0243	0.0521	-0.0350
---------	--------	--------	--------	---------	---------	---------	---------	--------	---------	--------	--------	---------

از اعداد فوق مشخص است که به ترتیب ستون‌های ۱ و ۳ ماتریس U همبستگی بیشتری با قیمت خانه دارند و می‌توان از آن‌ها برای بدست آوردن قیمت خانه استفاده کرد.

نمودار قیمت خانه بر حسب هر یک از این دو ویژگی جدید به صورت زیر خواهد بود:



همچنین نمودار قیمت خانه بر حسب این دو ویژگی جدید به صورت زیر خواهد بود:

