"بسمه تعالی"

گزارش پروژه پیاده سازی الگوریتم های FP\_Growth و Apriori

درس : داده کاوی

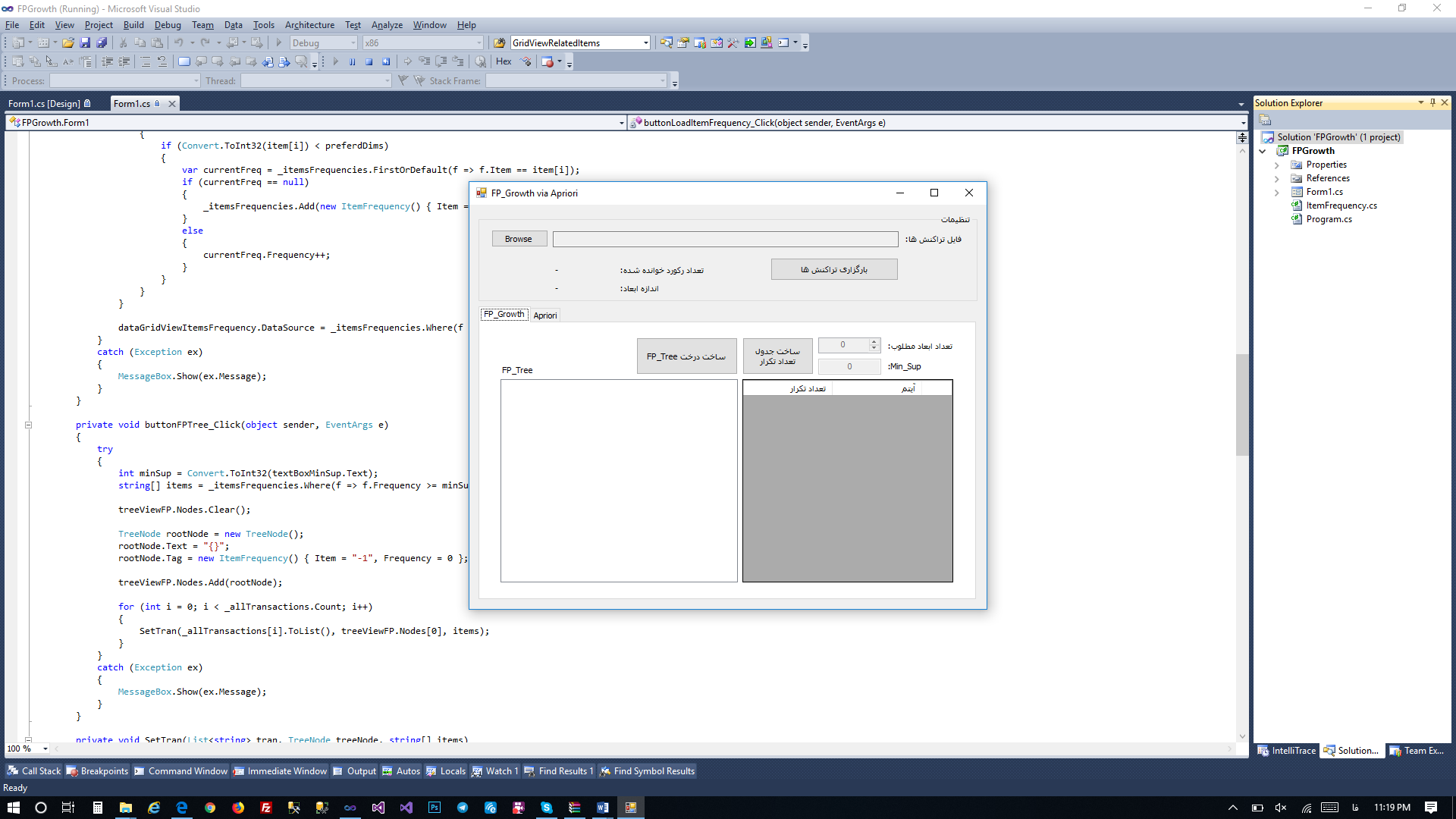
استاد : جناب آقای دکتر احمدی

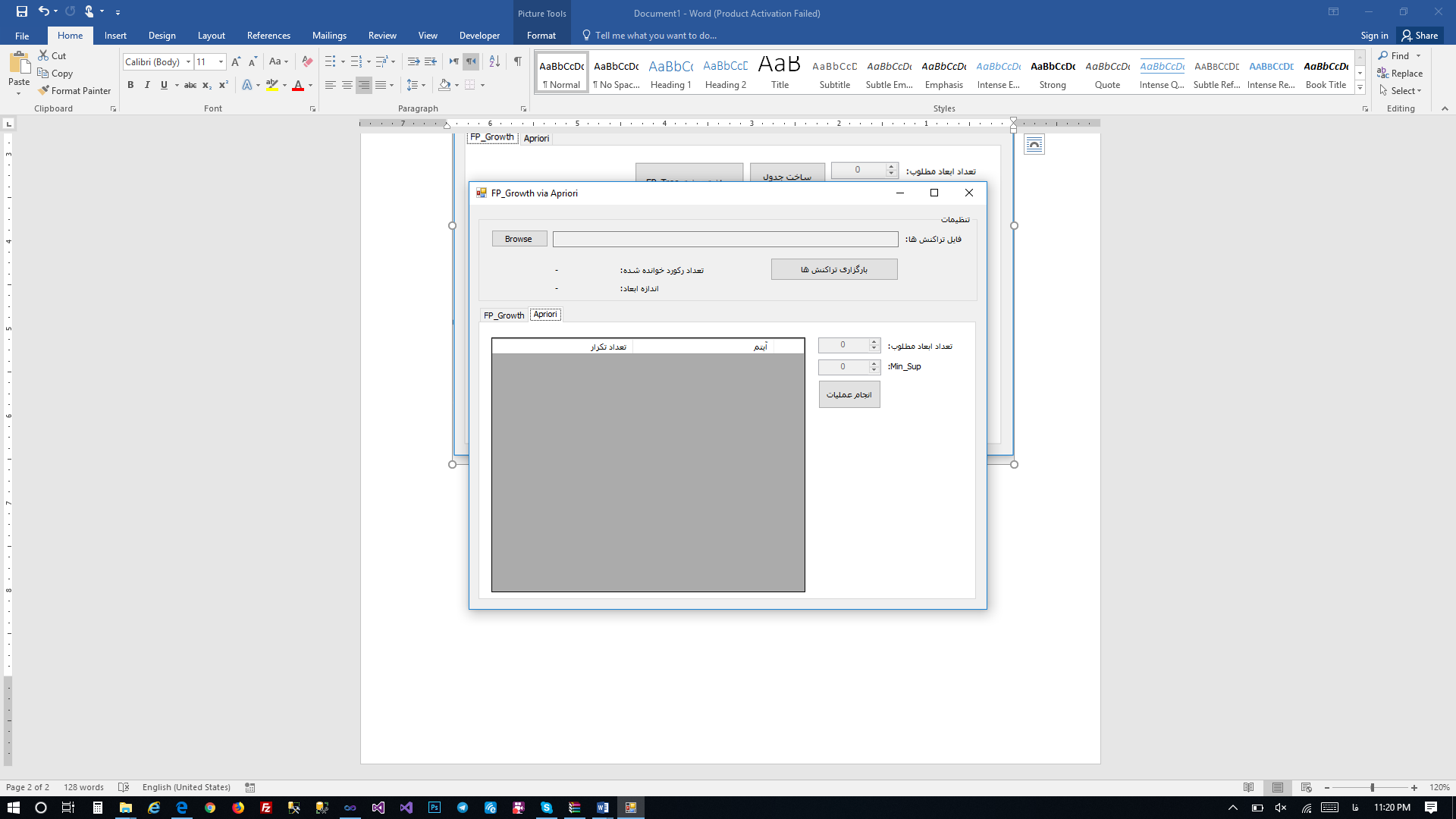
تهیه کننده : فرخنده زینالی آق قلعه

شماره دانشجویی : 9611274

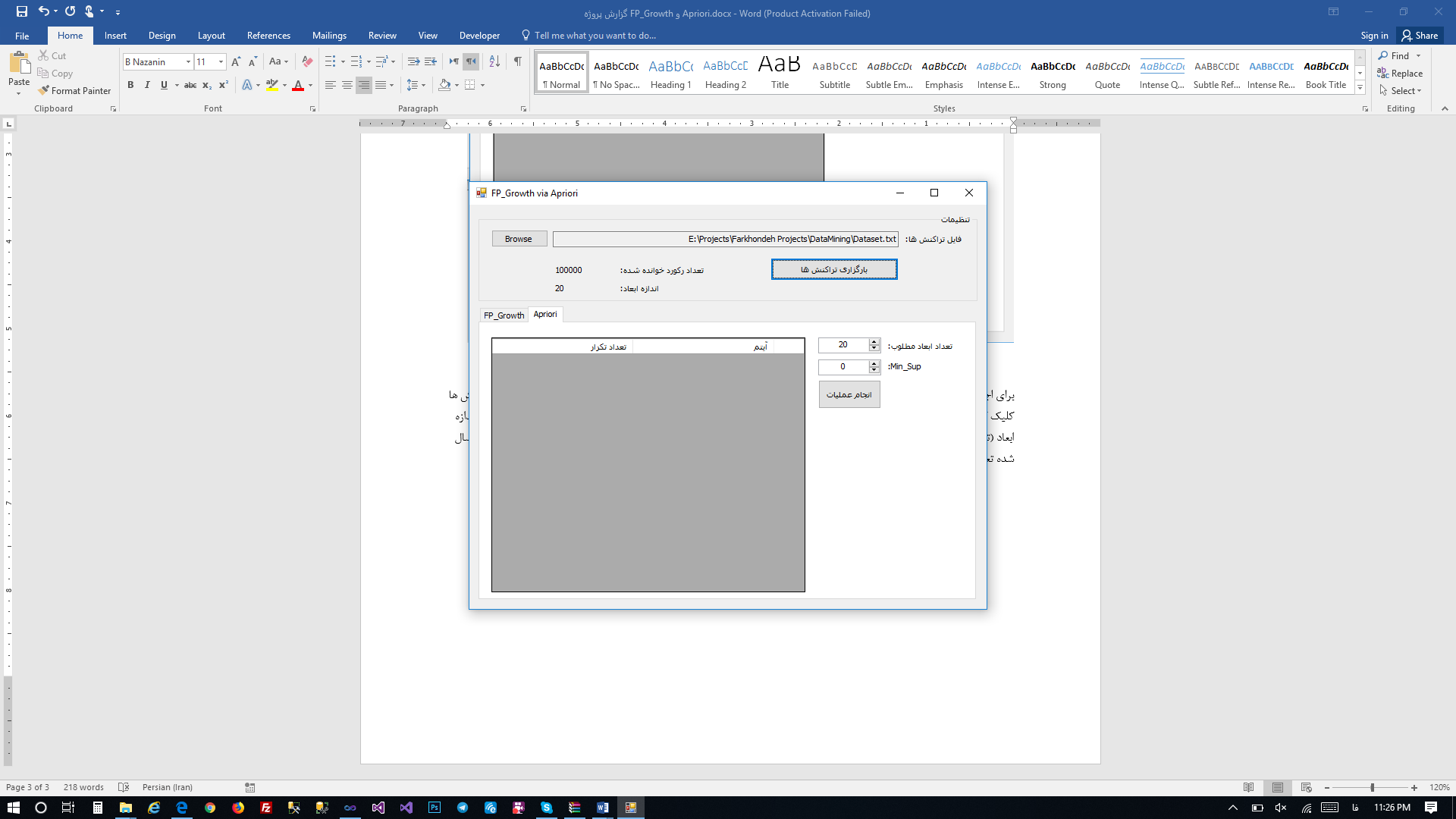
بهار 98

در این پروژه قصد داریم تا قسمتی از الگوریتم FP\_Growth یعنی ساخت درخت FP\_Tree و الگوریتم Apriori را پیاده سازی کنیم. برای انجام این کار پروژه ای از نوع Windows Application با زبان C# در محیط Visual Studio 2010 ایجاد کردیه ایم که نمای صفحه اصلی آن را در زیر مشاهده می نمایید. همانطور که ملاحظه می کنید قسمت مربوط به انتخاب فایل داده ها برای هر دو الگوریتم یکی می باشد امام جهت هر یک از الگوریتم های FP\_Growth و Apriori یک tab page جداگانه به همراه تنظیمات در نظر گرفته شده است.



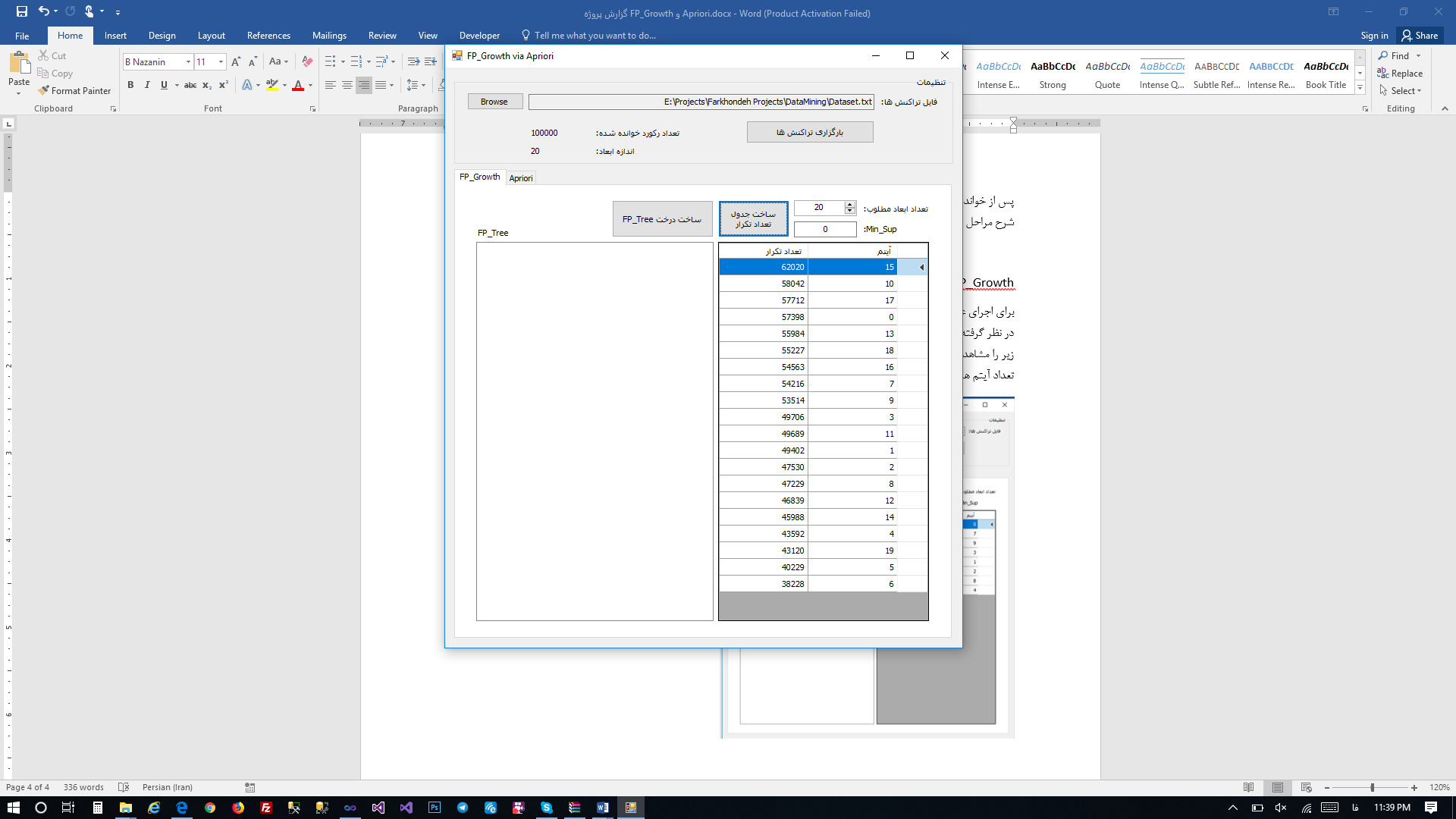
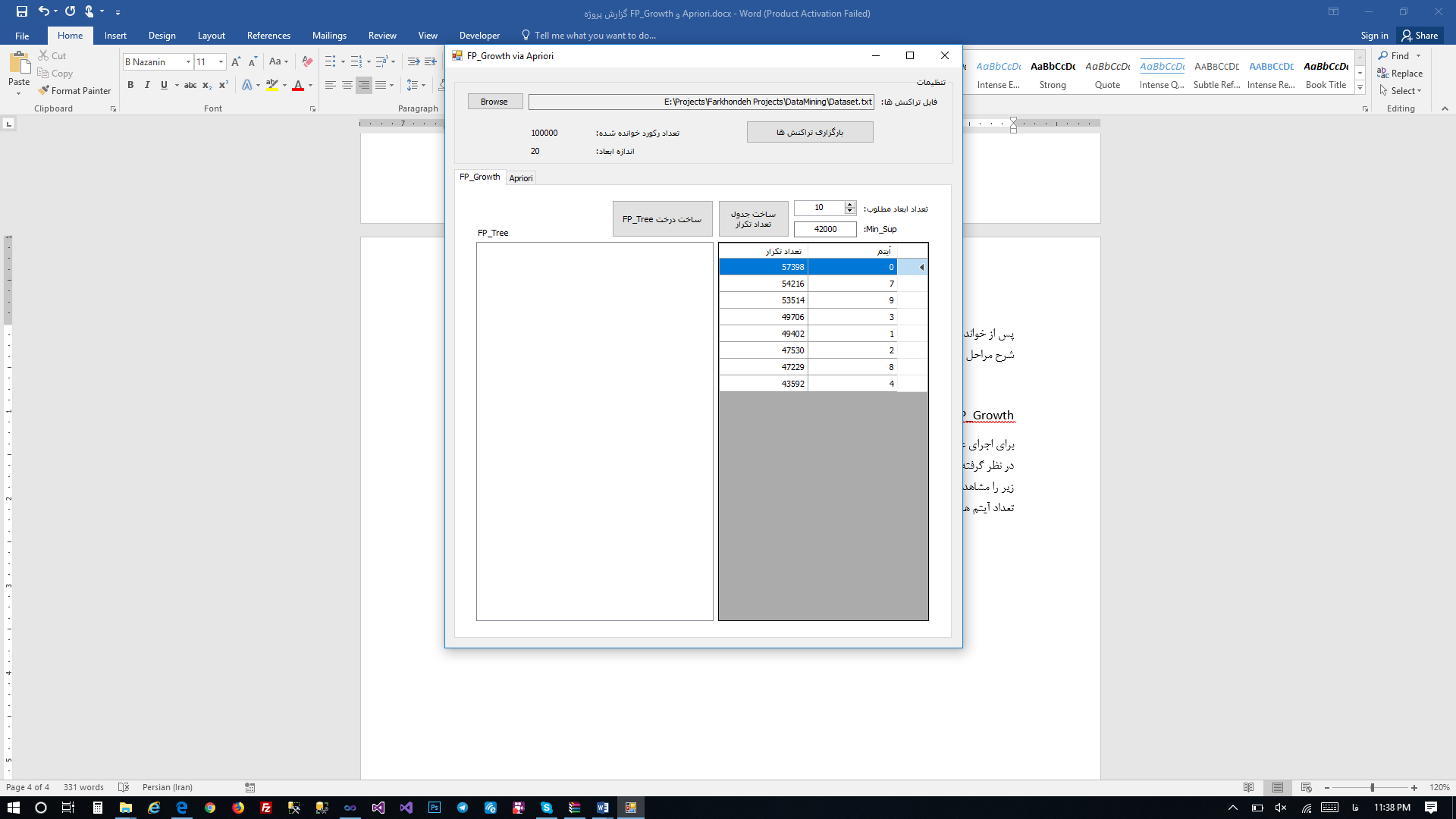


برای اجرای عملیات هر دوی الگوریتم ها ابتدا می بایست فایل داده ها را از قسمت تنظیمات انتخاب نمود و بر روی دکمه بارگزاری تراکنش ها کلیک کرد. در صورت موفقیت آمیز بودن خواندن فایل تراکنش ها، تعداد رکوردهای خوانده شده (تعداد ردیف های فایل تراکنش ها) و اندازه ابعاد (تعداد آیتم های متفاوت در تراکنش ها) نمایش داده می شود. همانطور که در تصویر زیر می بینید پس از انتخاب فایل داده های ارسال شده تعداد 100000 رکورد یا تراکنش و تعداد 20 آیتم یا بعد خوانده شده است.

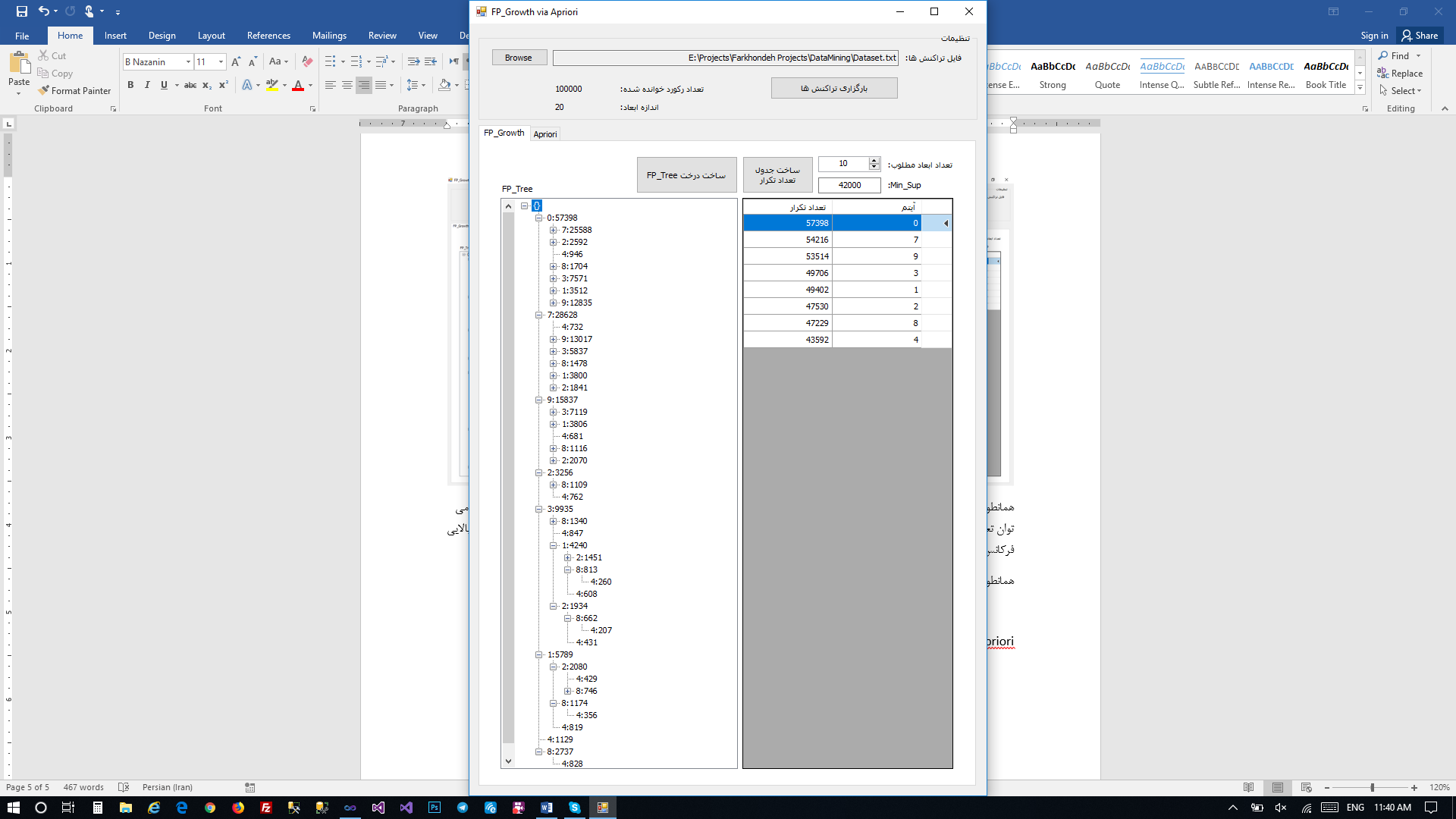


پس از خوانده شدن فایل تراکنش ها می توان یکی از عملیات FP\_Growth یا Apriori را از قسمت tab page پایین صفحه انتخاب و به شرح مراحل زیر اجرا نمود:

**FP\_Growth**

برای اجرای عملیات FP\_Growth ابتدا می بایست جدول تعداد تکرار آیتم ها را ایجاد نمود لذا تعداد ابعاد مطلوب یا همان تعداد آیتم های در نظر گرفته شده و همچنین مقدار Min\_Sup را وارد نموده و بر روی دکمه ساخت جدول تعداد تکرار کلیک نمایید. همانطور که دو تصویر زیر را مشاهده می کنید که یکی تعداد آیتم ها را 20 و Min\_Sup 0 انتخاب شده است و دیگری همانطور که صورت مسئله خواسته است تعداد آیتم های 10 و مقدار Min\_Sup 42000 را انتخاب نموده ایم.

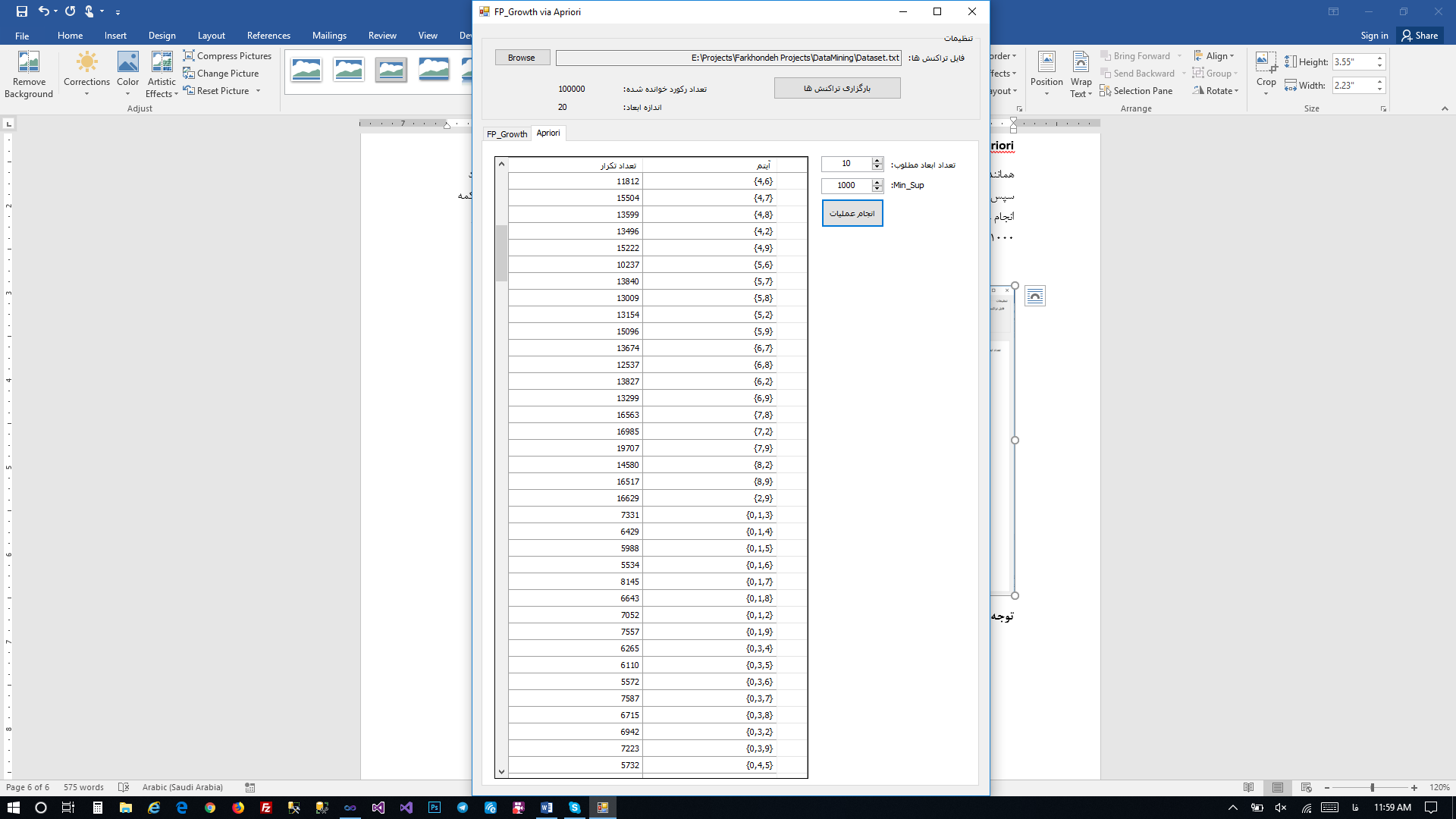
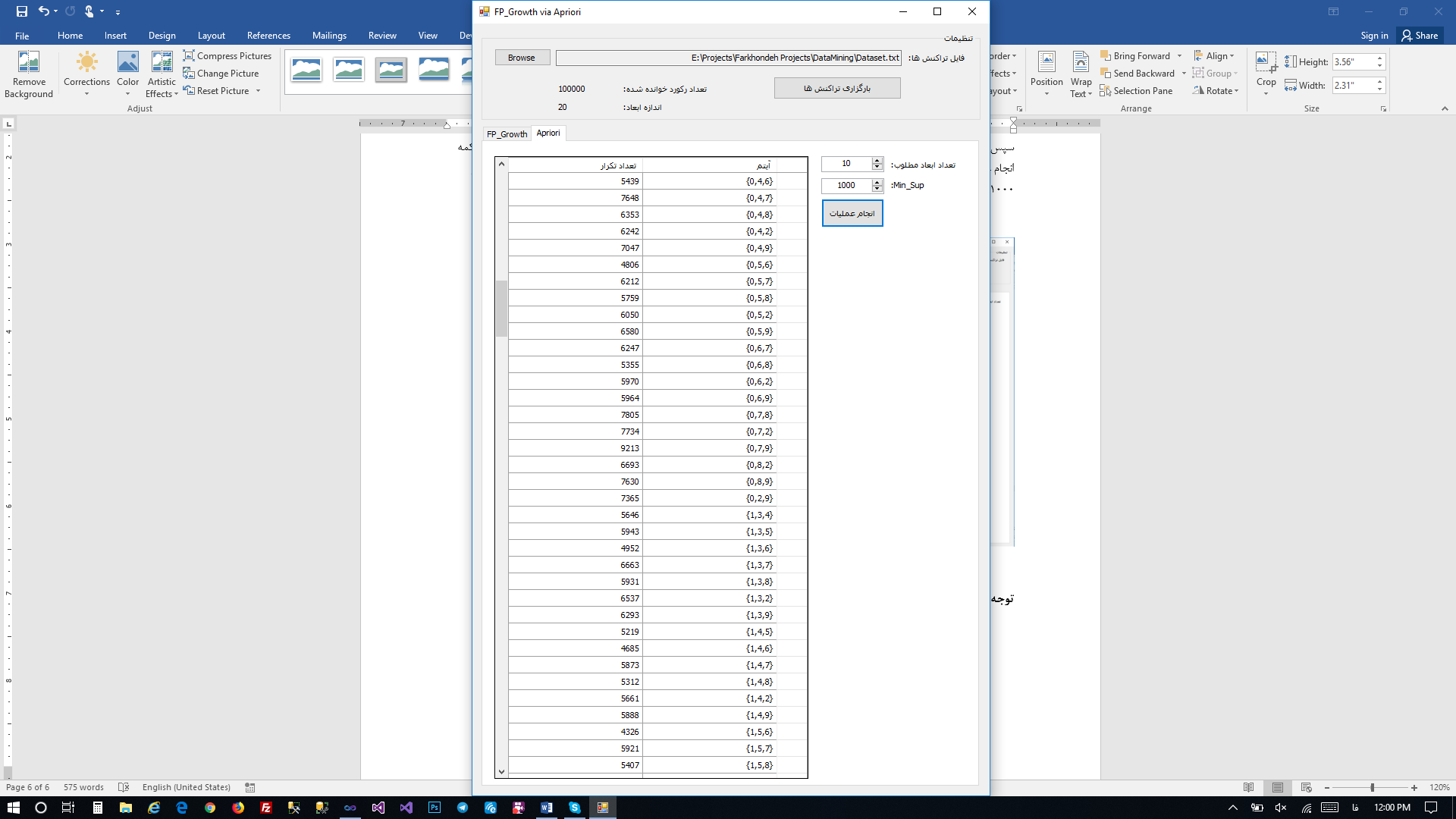
کار را با تنظیمات تصویر دوم یعنی تعدا ابعاد 10 و Min\_Sup 42000 ادامه می دهیم. برای ساخت درخت FP\_Tree بر روی دکمه ساخت درخت FP\_Tree کلیک می کنیم. پس از اتمام عملیات درخت مذکور به شکل زیر ساخته و نمایش داده می شود:

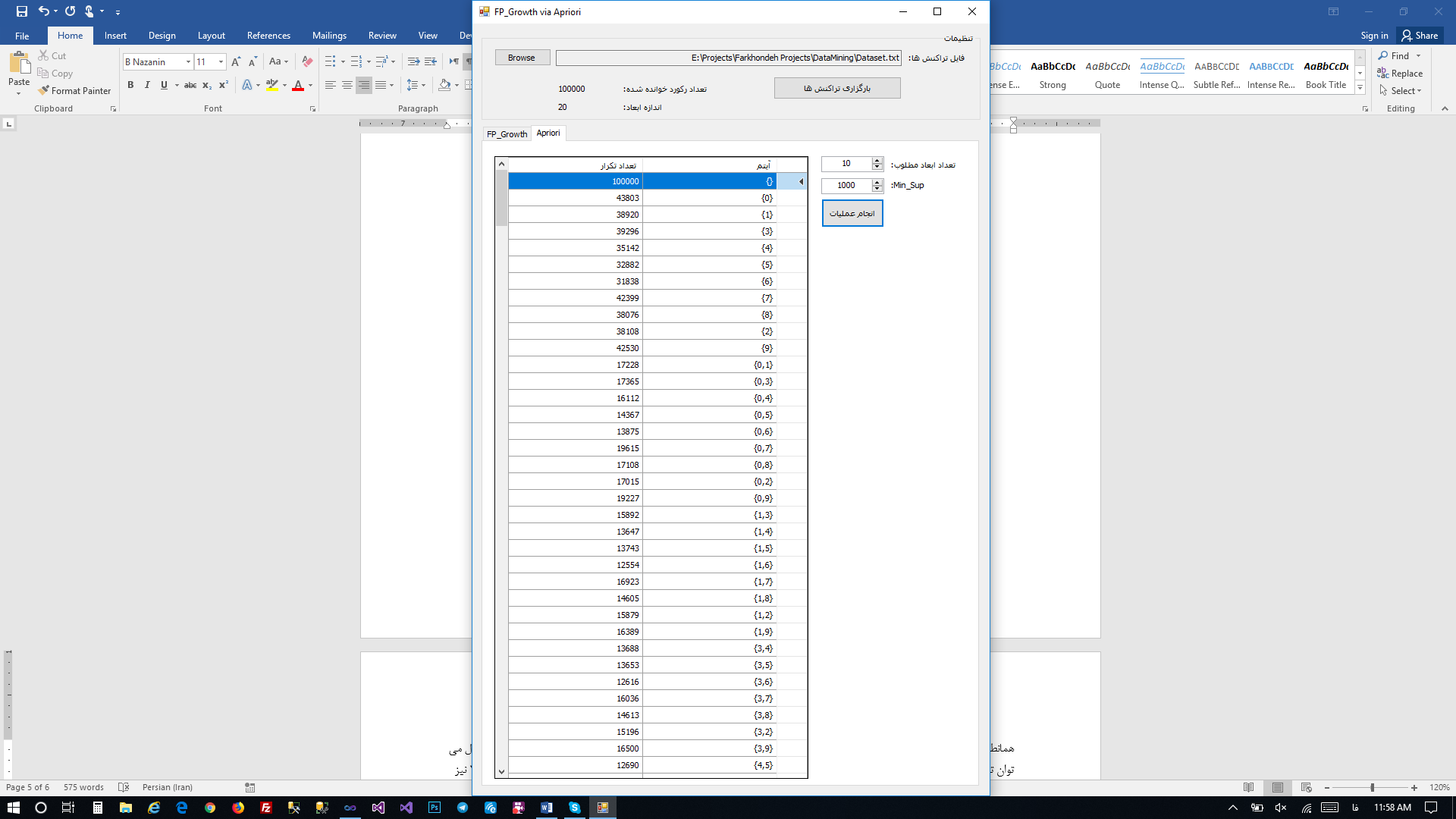


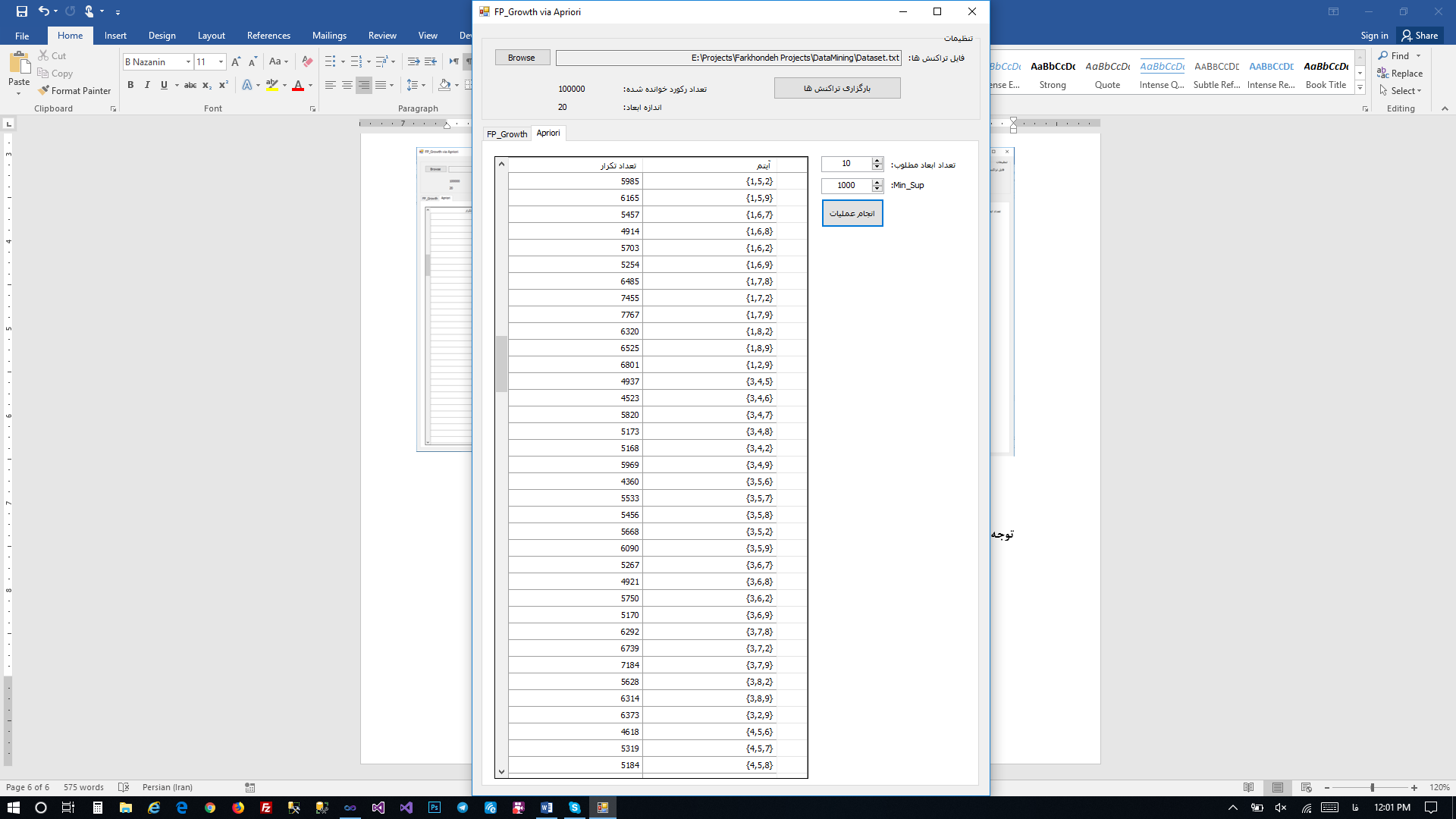
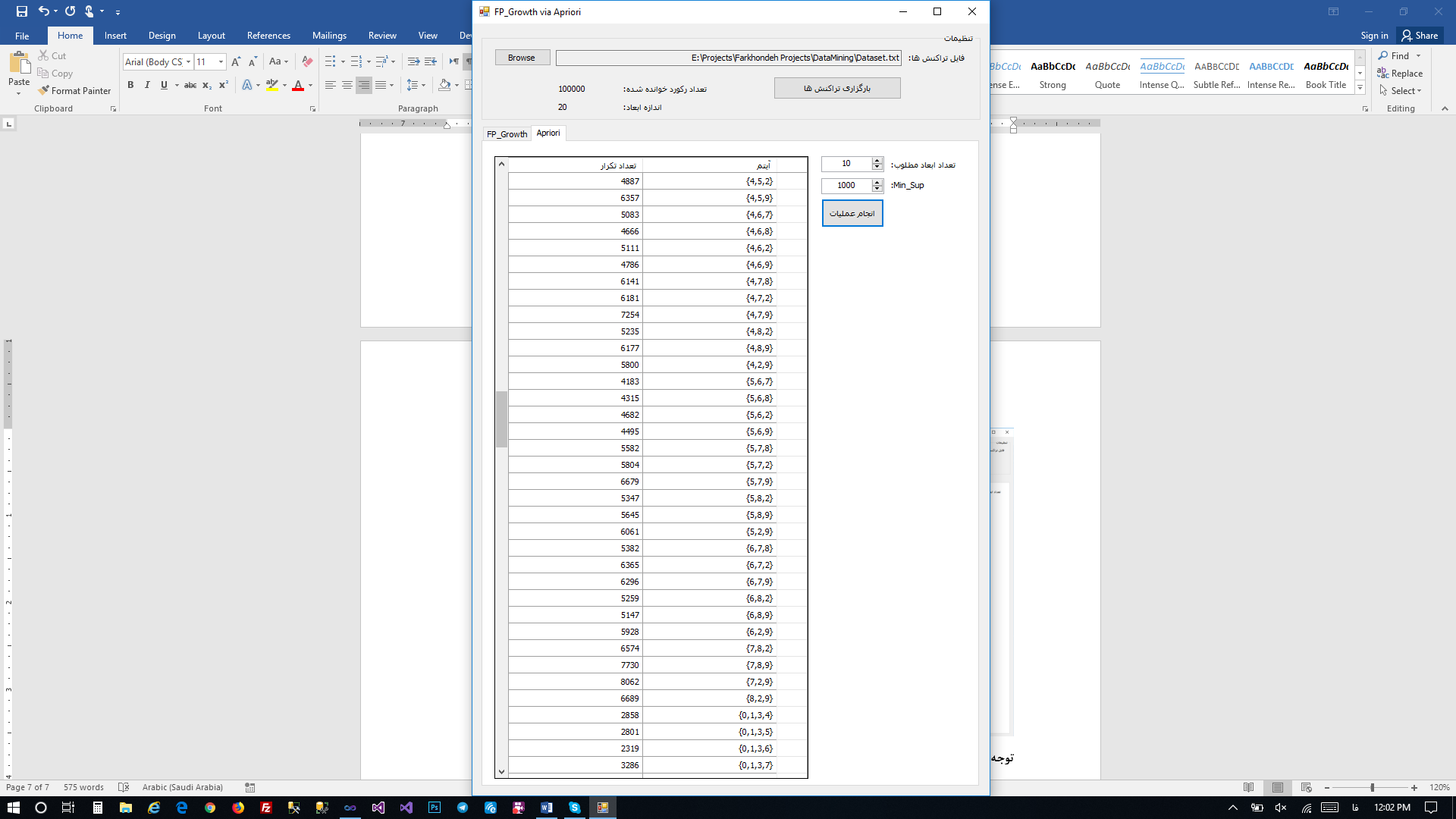
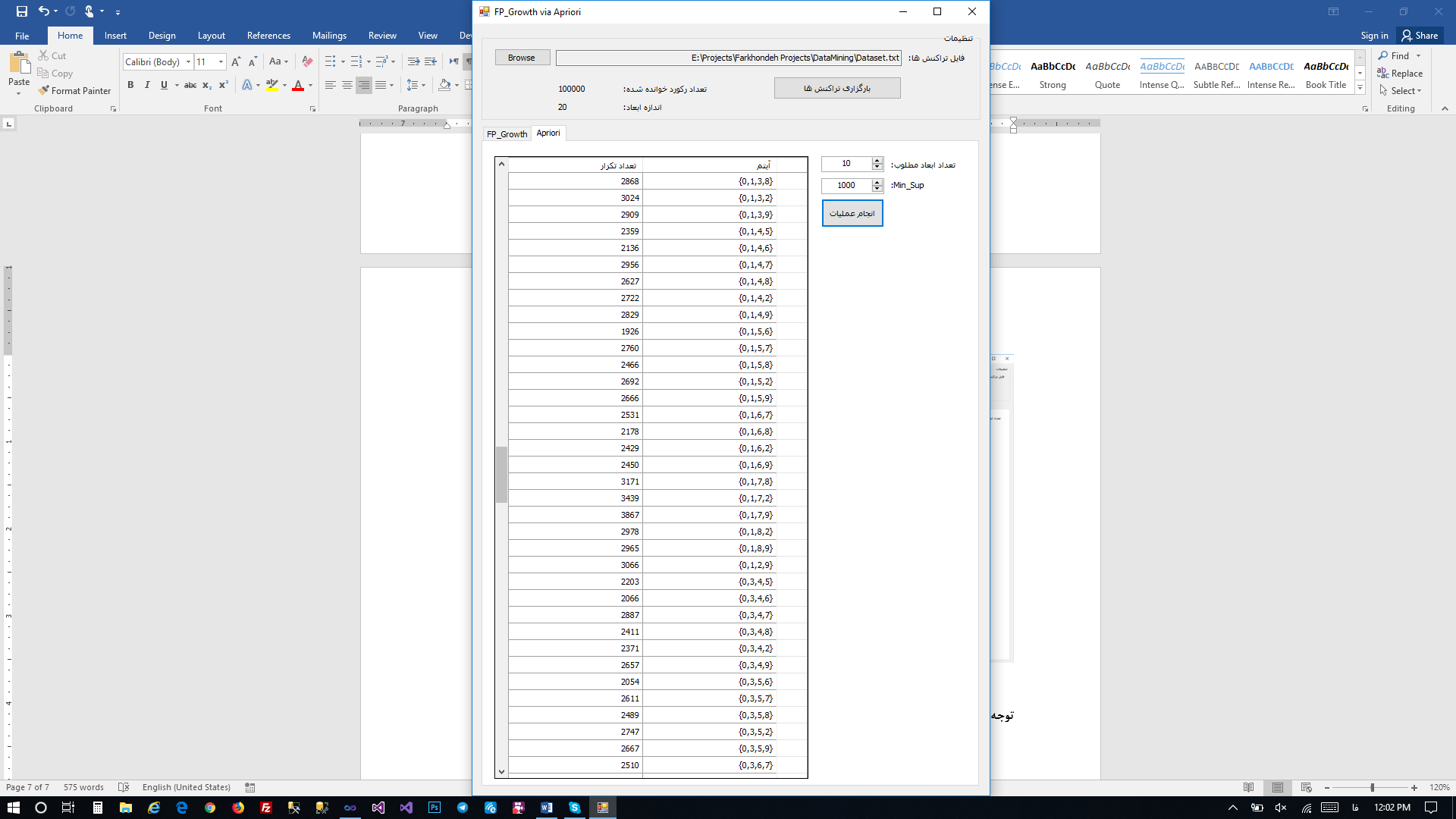
همانطور که مشاهده می کنید به علت کمبود جای نمایش نمی توان تمامی گره های درخت را در حالت باز نشان داد اما در همان نگاه اول می توان تعداد تکرار آیتم 0 را که بیشترین تکرار را دارد با فرکانس مربوطه در جدول تعداد تکرار و پس از آن آیتم بعدی یعنی آیتم شماره 7 نیز مجموع دو گره بالایی فرکانس آن را نشان می دهد و به همین صورت تا آخرین آیتم.

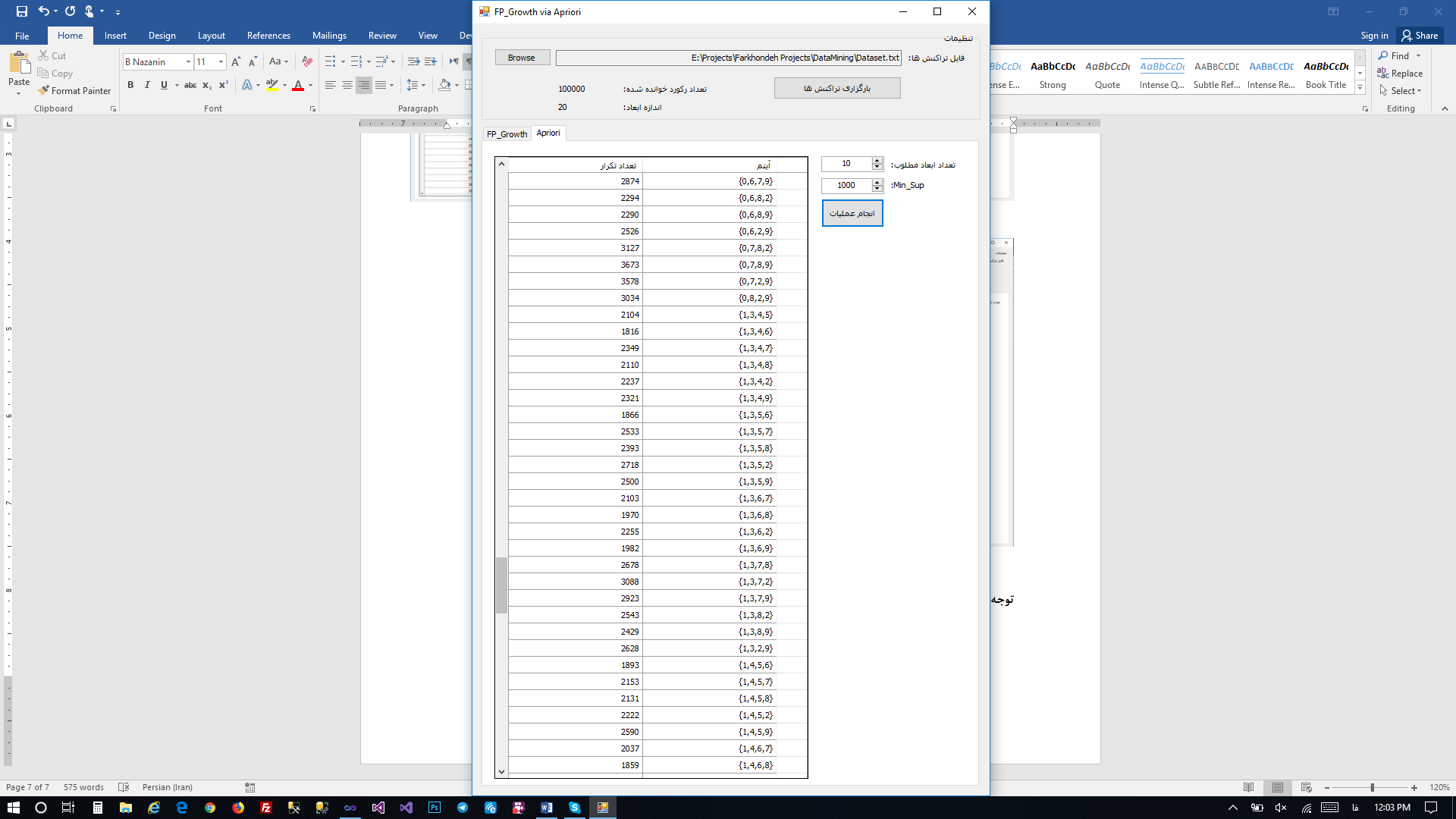
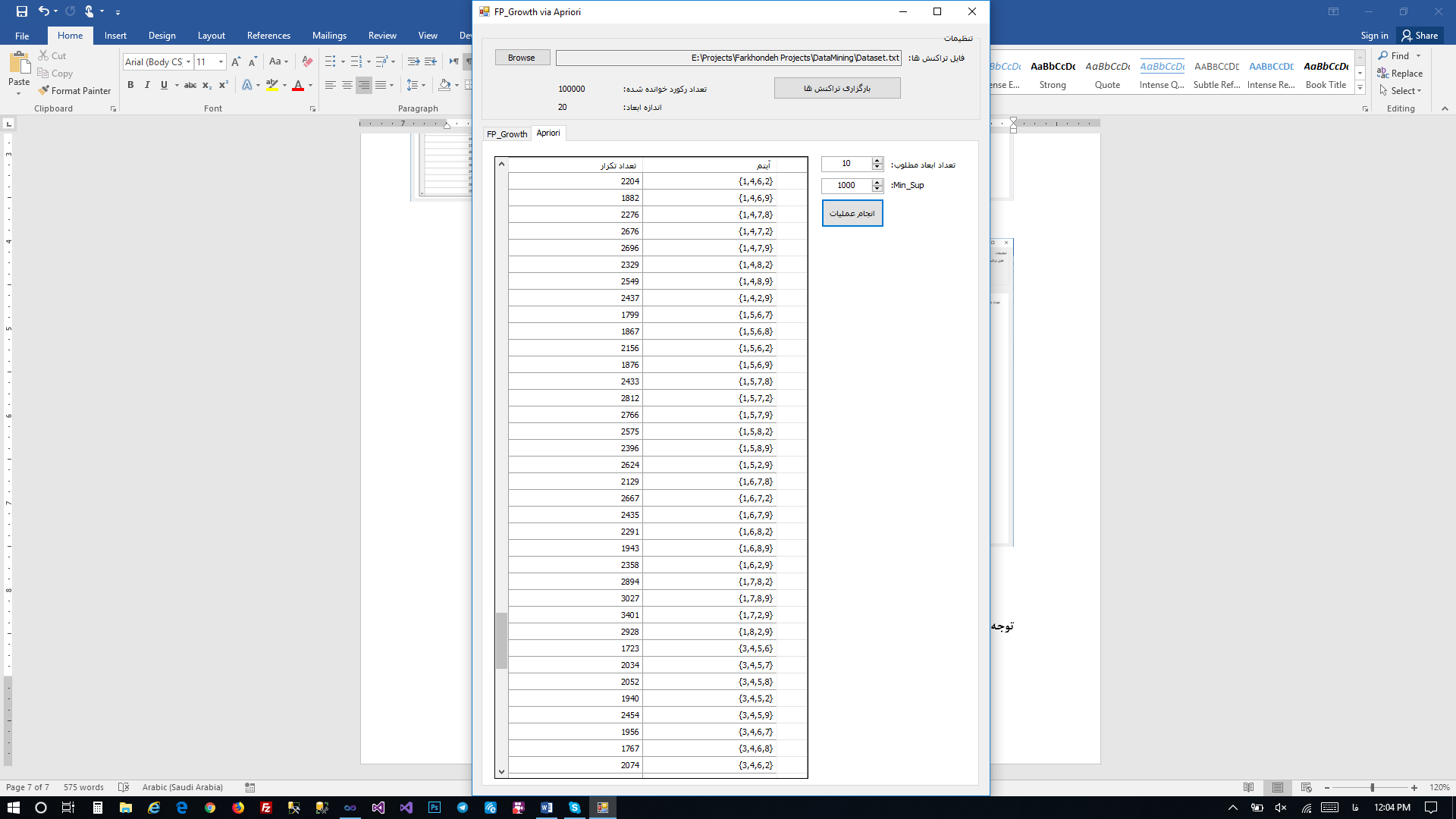
**Apriori**

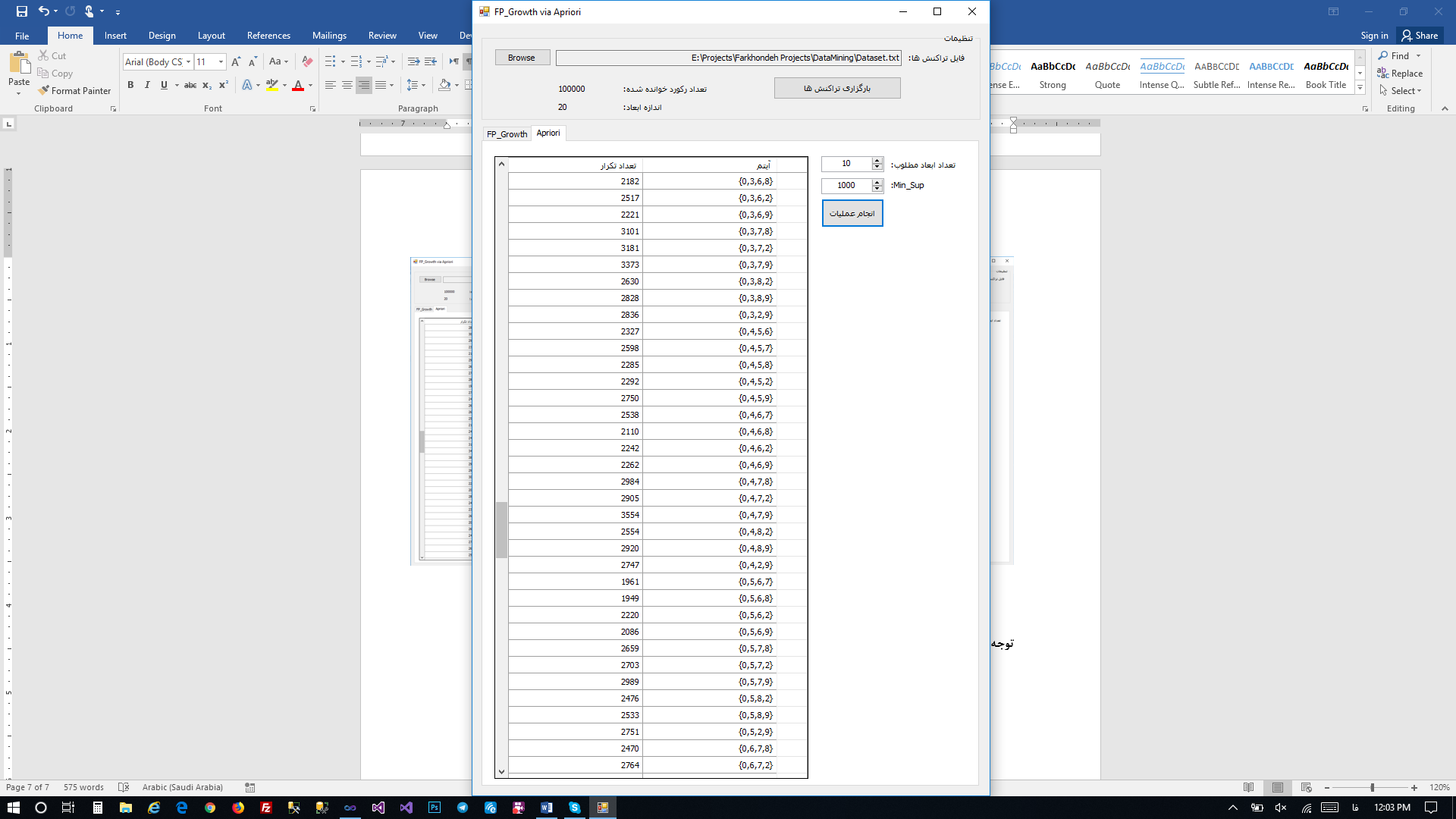
همانند بخش قبل فایل داده ها را انتخاب کرده و پس از بارگزاری تراکنش ها تعداد کل تراکنش ها و تعداد آیتم ها قابل مشاهده می شود سپس وارد tab page مربوط به Apriori می شویم و مقادیر تعداد ابعاد مطلوب و Min\_Sup را وارد می کنیم و پس از آن بر روی دکمه انجام عملیات کلیک می کنیم. همانطور که در صورت مسئله خواسته شده تعداد آیتم های مطلوب را برابر 10 و مثلا Min\_Sup را برابر 1000 قرار می دهیم و نتیجه تعداد تراکنش ها برای ترکیبات 1 تا 10 تایی که بیشتر از Min\_Sup هستند را در جدول می بینیم.

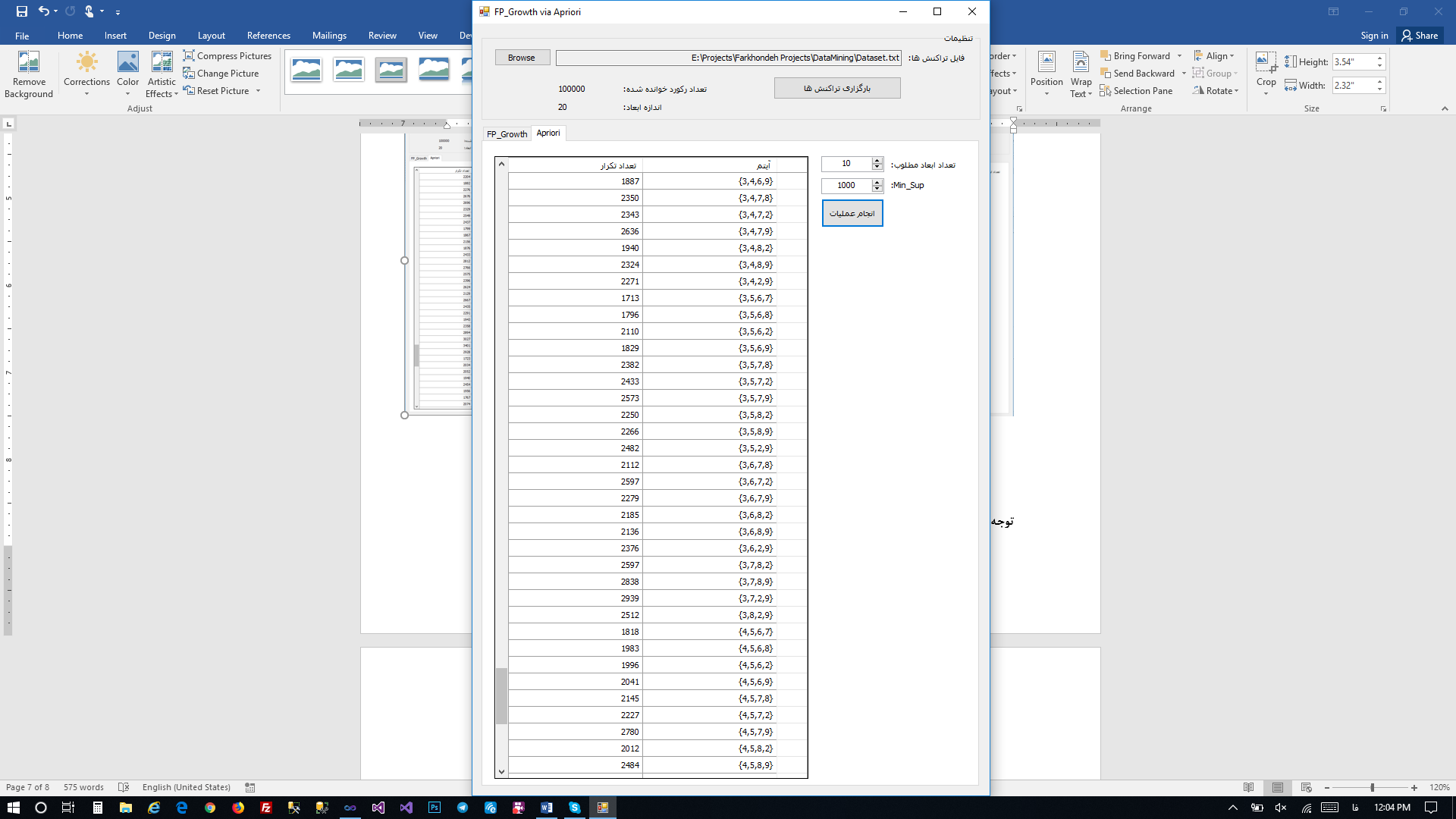
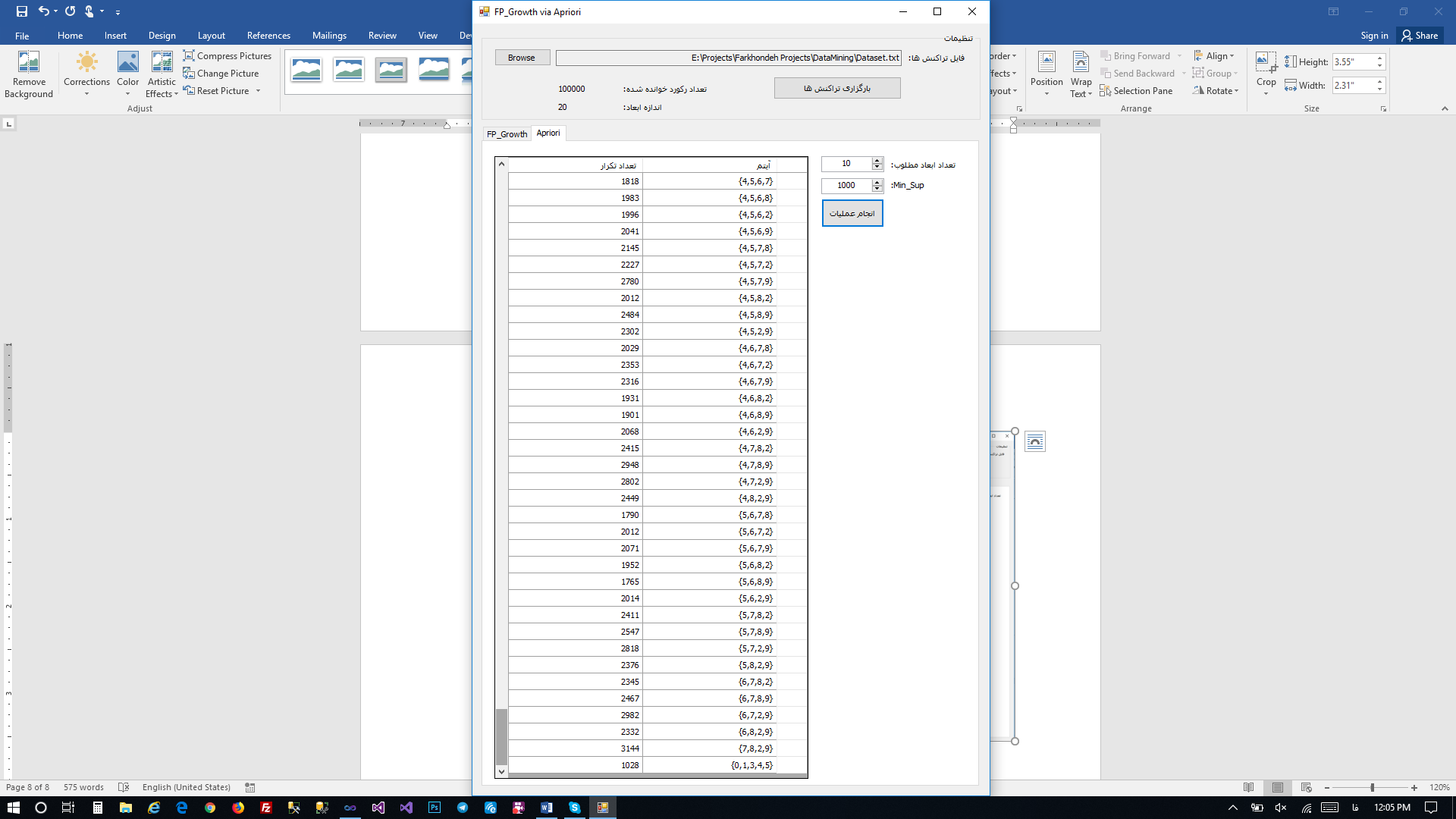












**توجه:** همانطور که مشاهده شد تعداد ابعاد مطلوب و همینطور Min\_Sup در پروژه قابل تنظیم می باشد.