In [1]:

```
import pandas as pd
import numpy as np
import cufflinks as cf
from plotly.offline import download_plotlyjs, init_notebook_mode, plot, iplot
init_notebook_mode(connected = True) # to get the connection
cf.go_offline()#offline mode
from sklearn.model_selection import train_test_split
from scipy.sparse import hstack,csr_matrix
from sklearn.model_selection import GridSearchCV
from sklearn.linear_model import SGDClassifier
import warnings
warnings.filterwarnings("ignore")
from scipy.sparse import hstack,vstack,csr_matrix
from sklearn.model_selection import PredefinedSplit
from sklearn.metrics import roc_auc_score,confusion_matrix,roc_curve, auc
import matplotlib.pyplot as plt
from sklearn.feature_selection import SelectKBest, chi2
from vaderSentiment.vaderSentiment import SentimentIntensityAnalyzer
from sklearn.preprocessing import StandardScaler,Normalizer
import scipy
from tqdm import tqdm
from sklearn.cluster import KMeans,AgglomerativeClustering,DBSCAN
from sklearn.decomposition import TruncatedSVD
%matplotlib inline
```

1.Reading and Splitting Data

In [2]:

```
#considering 10k points only due to constraints
data=pd.read_csv('/home/sathish/Desktop/appliedai/Donors_choose_preprocessed_data.csv',na_filter=False)
data=data.sample(10000,random_state=42)
x_train=data.drop(['project_is_approved'],axis=1)
y=data['project_is_approved']
```

2. Creating Data Matrix

2.1 converting text to vectors (Bow)

In [3]:

```
from sklearn.feature_extraction.text import CountVectorizer
vectorizer=CountVectorizer()
x_train_essay_bow=vectorizer.fit_transform(x_train['essay'].values)
x_train_summary_bow=vectorizer.fit_transform(x_train['project_resource_summary'].values)
x_train_title_bow=vectorizer.fit_transform(x_train['project_title'].values)
```

2.2 one hot encoding the categorical features

In [4]:

```
vectorizer=CountVectorizer()
x_train_state_onehot=vectorizer.fit_transform(x_train['school_state'].values)
x_train_tprefix_onehot=vectorizer.fit_transform(x_train['teacher_prefix'].values)
x_train_grade_onehot=vectorizer.fit_transform(x_train['project_grade_category'].values)
x_train_cat_onehot=vectorizer.fit_transform(x_train['project_subject_categories'].values)
x_train_subcat_onehot=vectorizer.fit_transform(x_train['project_subject_subcategories'].values)
```

2.3 Normalizing the numerical features

In [5]:

```
normalizer=Normalizer()
```

```
x_train_price_norm=normalizer.fit_transform(x_train['price'].values.reshape(-1,1))
x_train_quantity_norm=normalizer.fit_transform(x_train['quantity'].values.reshape(-1,1))
x_train_prev_projects_norm=normalizer.fit_transform(x_train['teacher_number_of_previously_posted_projects'].values.reshape(-1,1))
```

2.4 Building data matrix

In [6]:

```
#text features
bow_train=hstack((x_train_essay_bow,x_train_summary_bow,x_train_title_bow))


#all categrocial features
x_train_cat=hstack((x_train_state_onehot,x_train_tprefix_onehot,x_train_grade_onehot,x_train_cat_onehot,x_train_subcat_onehot))

#numerical features
x_train_num=hstack((csr_matrix(x_train_price_norm),x_train_quantity_norm,x_train_prev_projects_norm))

#final matrices
final_bow_train=hstack((bow_train,x_train_cat,x_train_num)).tocsr()
```

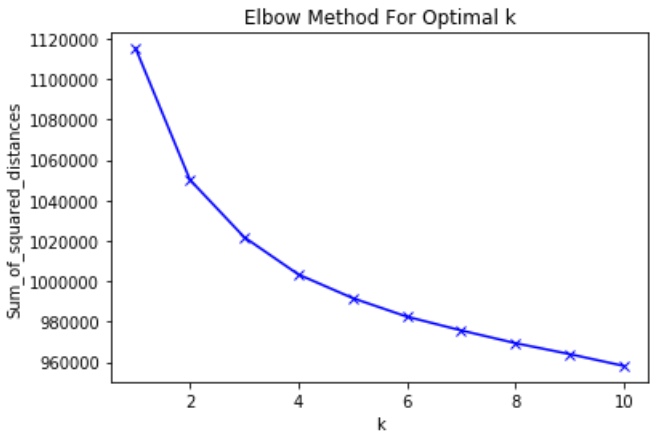
In [7]:

```
best_5k=SelectKBest(chi2,k=5000)
x_train_new=best_5k.fit_transform(final_bow_train,y)
```

3. Applying Kmeans

In [8]:

```
#https://blog.cambridgespark.com/how-to-determine-the-optimal-number-of-clusters-for-k-means-clustering-14f27070048f
Sum_of_squared_distances = []
K = range(1,11)
for k in K:
    km = KMeans(n_clusters=k)
    km = km.fit(x_train_new)
    Sum_of_squared_distances.append(km.inertia_)
plt.plot(K, Sum_of_squared_distances, 'bx-')
plt.xlabel('k')
plt.ylabel('Sum_of_squared_distances')
plt.title('Elbow Method For Optimal k')
plt.show()
```



In [9]:

```
#optimal k=5
kmeans=KMeans(n_clusters=5)
kmeans.fit(x_train_new)
predicted=kmeans.predict(x_train_new)
```

In [10]:

```
from wordcloud import WordCloud, STOPWORDS
def wordcloud(temp):
```

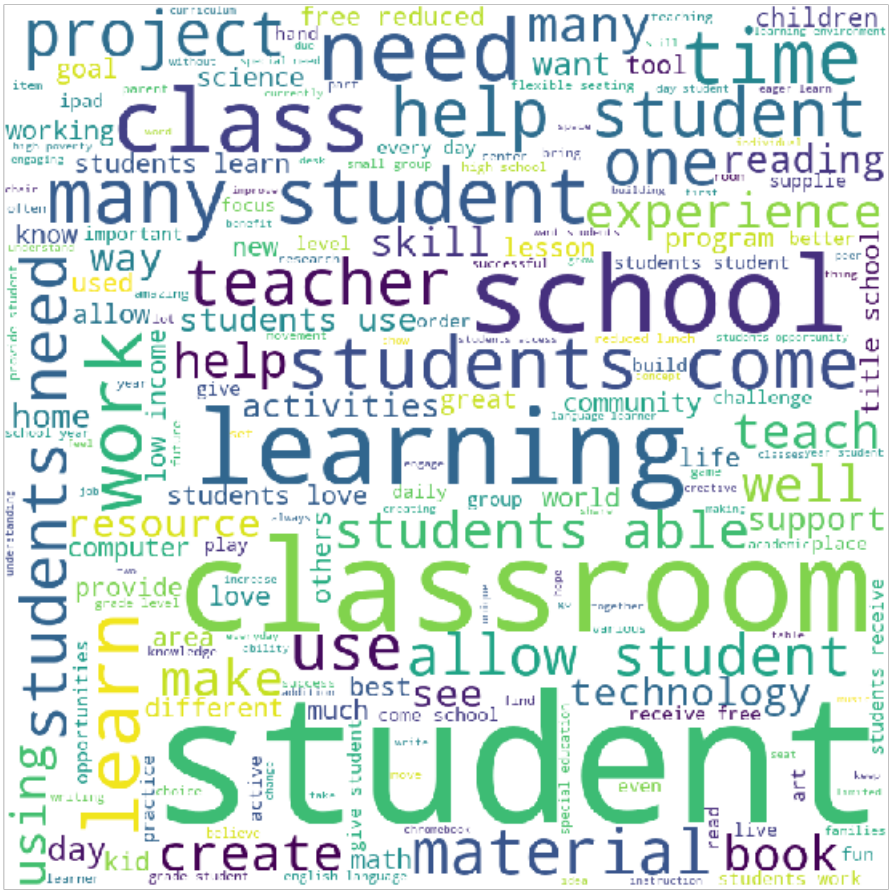
```
def wordcloud(temp):
    values=data[temp]['essay']
    comment_words = ' '
    stopwords = set(STOPWORDS)

    for val in values:

        # typecaste each val to string
        val = str(val)

        # split the value
        tokens = val.split()

        # Converts each token into lowercase
        for i in range(len(tokens)):
            tokens[i] = tokens[i].lower()

        for words in tokens:
            comment_words = comment_words + words + ' '


    wordcloud = WordCloud(width = 800, height = 800,
                    background_color ='white',
                    stopwords = stopwords,
                    min_font_size = 10).generate(comment_words)

    # plot the WordCloud image
    plt.figure(figsize = (8, 8), facecolor = None)
    plt.imshow(wordcloud)
    plt.axis("off")
    plt.tight_layout(pad = 0)

    plt.show()
```

In [11]:

```
clusters=5
for i in range(clusters):
    print(("*"*20)+str(i)+" cluster"+("*"*20))
    temp=(predicted==i)
    print(data[temp][:5])
    wordcloud(temp)
```

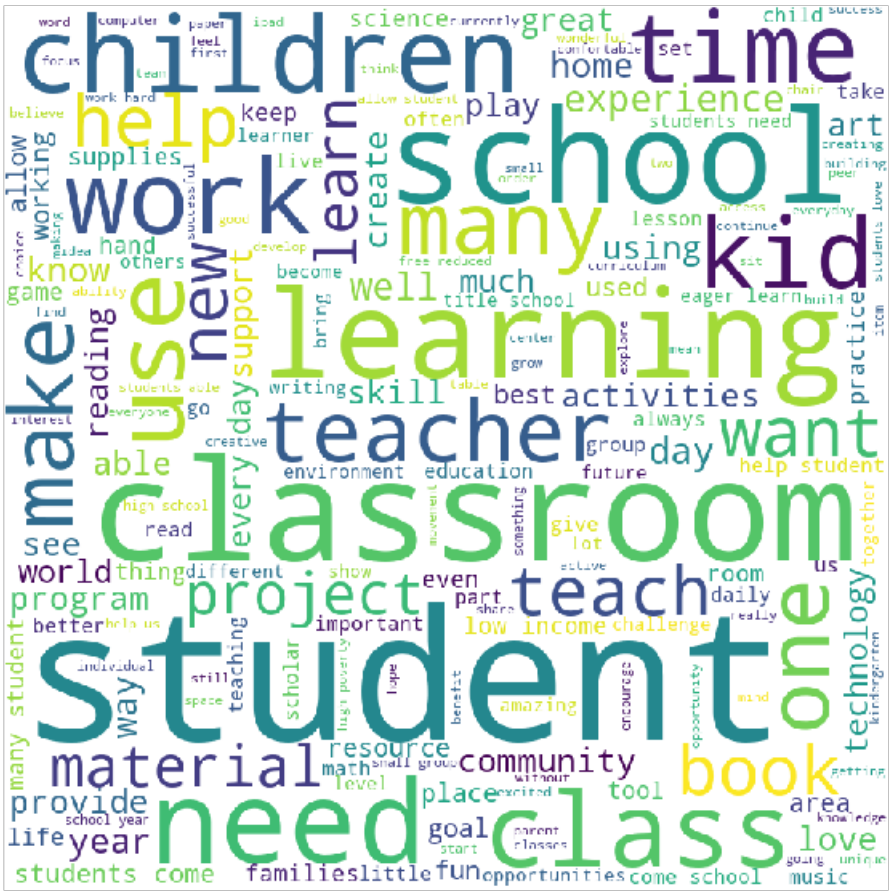
```
********************0 cluster********************
            id teacher_prefix school_state project_submitted_datetime  \
102804  p036514           Mrs.           CA        2016-06-29 01:40:00
100478  p140242            Ms.           CT        2016-09-19 18:04:30
42921   p184874        Teacher           IL        2016-10-26 21:13:32
83483   p153579           Mr.            SC        2016-08-12 08:17:47
13814   p171057            Ms.           TN        2016-11-14 23:09:12

       project_grade_category             project_subject_categories  \
102804         Grades_PreK_2    Literacy & Language, Math & Science
100478         Grades_PreK_2     Literacy & Language, Special Needs
42921             Grades_6_8    Math & Science, Literacy & Language
83483             Grades_3_5  Literacy & Language, Music & The Arts
13814             Grades_3_5  History & Civics, Literacy & Language

                    project_subject_subcategories  \
102804                         Literacy, Mathematics
100478                        Literacy, Special Needs
42921   Health & Life Science, Literature & Writing
83483                                 Literacy, Music
13814     History & Geography, Literature & Writing

                                        project_title  \
102804                   forget paper let go eco friendly
100478  please help my students with autism practice t...
42921           flipping our classroom into 21st century
83483                                 minds with mallets
13814                         working through our wiggles

                              project_resource_summary  \
102804  my students need boogie boards inexpensive reu...
100478  my students need learning tools reinforce read...
42921   my students need chromebooks flip classroom 21...
```

```
83483   my students need increase number orff instrume...
13814   my students need chairs wiggle safely work des...

        teacher_number_of_previously_posted_projects  project_is_approved  \
102804                                            10                   1
100478                                             2                   0
42921                                              3                   1
83483                                              1                   0
13814                                              1                   1

                                                    essay   price  quantity
102804  i classroom wonderful students we start learni...   29.99         4
100478  i teach classroom 8 students autism a typical ...  269.95         6
42921   as teacher low income high poverty school dist...  179.00         4
83483   i teach music 600 students school i love worki...  955.50         4
13814   our school awesome community school students t...  220.46        15
```



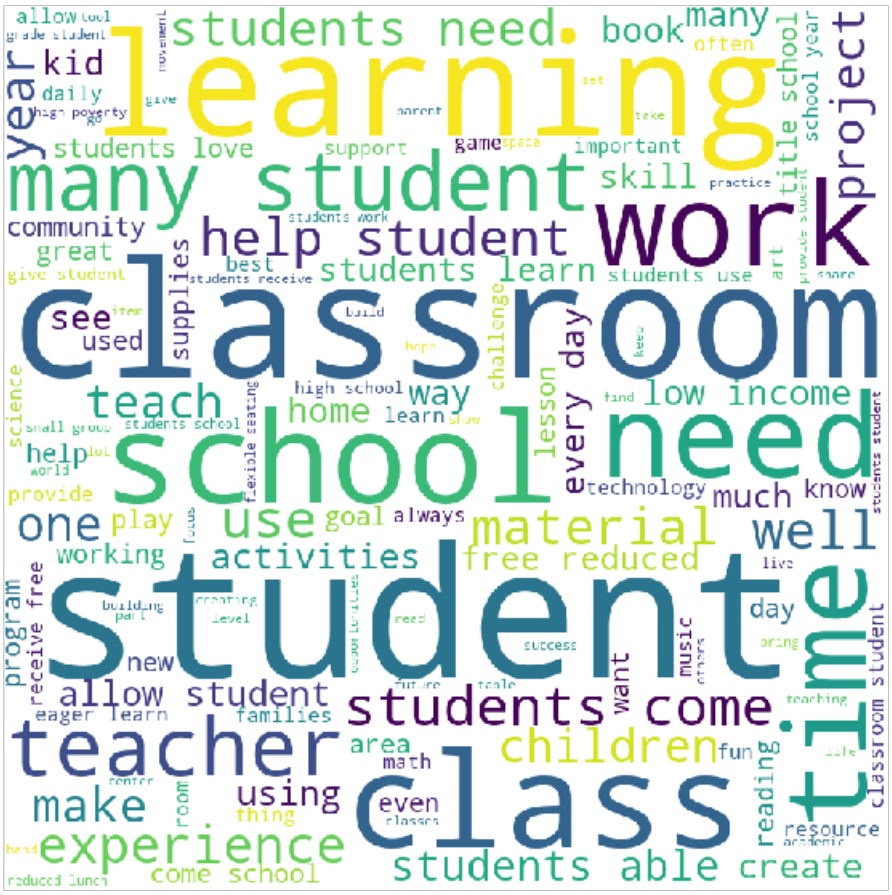

```
********************1 cluster********************
            id teacher_prefix school_state project_submitted_datetime  \
59452  p226149           Mrs.           ME        2017-03-07 13:01:09
24971  p254758            Mr.           MD        2016-06-20 10:14:16
47714  p149892           Mrs.           NY        2016-10-25 19:38:43
10776  p062274            Ms.           KY        2017-03-28 18:22:57
48607  p096186           Mrs.           MA        2017-01-04 14:36:29

      project_grade_category           project_subject_categories  \
59452          Grades_PreK_2                Literacy & Language
24971             Grades_3_5                     Math & Science
47714             Grades_3_5                Literacy & Language
10776          Grades_PreK_2  Literacy & Language, Math & Science
48607             Grades_6_8                     Math & Science

                project_subject_subcategories             project_title  \
59452                             Literacy   spruce seuss collection
24971                Environmental Science        how plants help us
47714                 Literature & Writing       as we grow we learn
10776                Literacy, Mathematics          wobbling working
48607  Applied Sciences, Environmental Science  do not throw away nation

                                project_resource_summary  \
59452  my students need dr seuss books add learning l...
24971  my students need lab equipment test plants imp...
```

```
47714  my students need working printer access additi...
10776  my students need 5 wobble stools replace yoga ...
48607  my students need asus zen phones vr googles cr...

       teacher_number_of_previously_posted_projects  project_is_approved  \
59452                                            21                    1
24971                                            12                    1
47714                                             3                    1
10776                                             5                    1
48607                                            14                    0

                                                   essay   price  quantity
59452  my students come various backgrounds we teach ...  192.94        28
24971  we fabulous group fifth grade students baltimo...  389.64         3
47714  i lucky year boisterous talented group 28 thir...  233.08         2
10776  our classroom helps get wiggles i work this co...   50.12         5
48607  nrms located cape cod our local area supported...  387.08        23
```



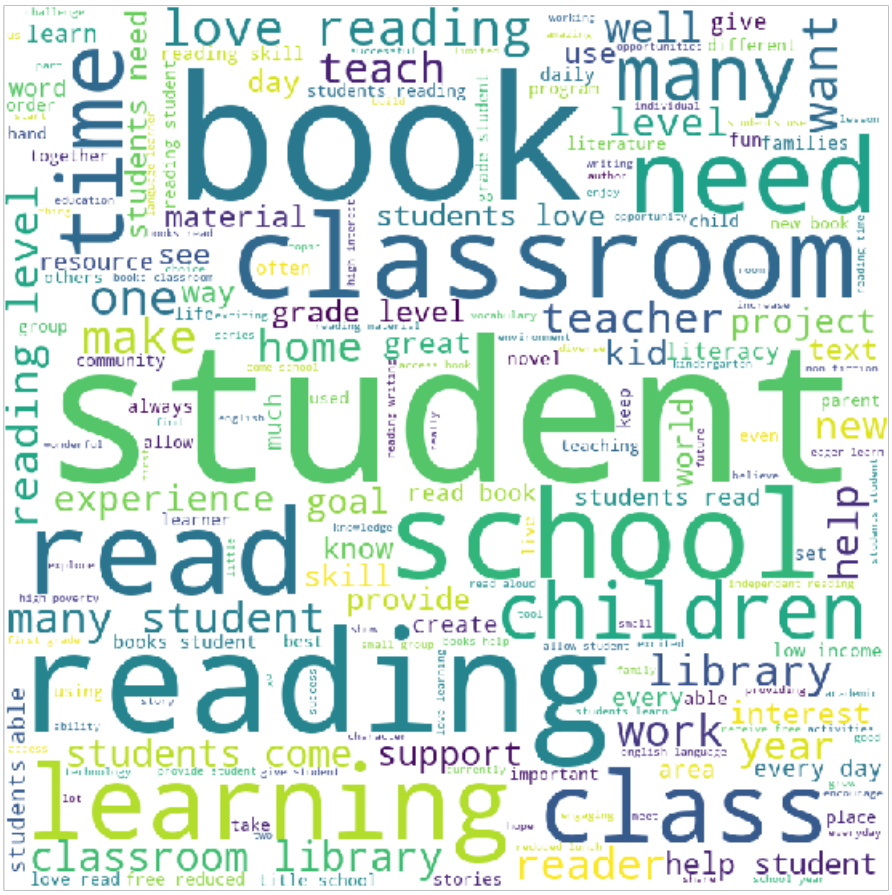

```
**********2 cluster**********
            id teacher_prefix school_state project_submitted_datetime  \
105447  p118457        Teacher           NC        2016-07-05 16:12:10
69262   p013133           Mrs.           NY        2016-10-15 16:53:46
55966   p205746            Ms.           UT        2016-08-25 01:29:34
35390   p158955            Ms.           CA        2016-07-22 16:42:42
88354   p047488            Ms.           HI        2016-07-01 17:49:36

       project_grade_category     project_subject_categories  \
105447           Grades_9_12  Math & Science, Special Needs
69262             Grades_3_5                Special Needs
55966            Grades_9_12              History & Civics
35390             Grades_6_8                Math & Science
88354             Grades_6_8                Math & Science

                project_subject_subcategories  \
105447     Environmental Science, Special Needs
69262                           Special Needs
55966      History & Geography, Social Sciences
35390                   Environmental Science
88354   Applied Sciences, Environmental Science

                                project_title  \
105447                 rain barrels school garden
```

```
69262                              wobbly learning
55966                hold tight using ipad case grip
35390                                  click engage
88354   life science lab materials let investigate

                                project_resource_summary  \
105447  my students need rain barrels collect water ga...
69262   my students need five wobble chair they energy...
55966   my students need ipad case drop new ipad not b...
35390   my students need 34 interactive clickers make ...
88354   my students need materials conduct biology lab...

        teacher_number_of_previously_posted_projects  project_is_approved  \
105447                                             0                    1
69262                                             20                    0
55966                                              2                    1
35390                                              0                    1
88354                                              9                    1

                                                    essay    price  quantity
105447  environmental science students no idea food co...    93.69         4
69262   my students need six learniture active learnin...   209.85         5
55966   our high school school residing salt lake city...   102.31         2
35390   my middle school located underserved community...  1541.14         3
88354   my students awesome they continually amaze fre...   188.53         8
```



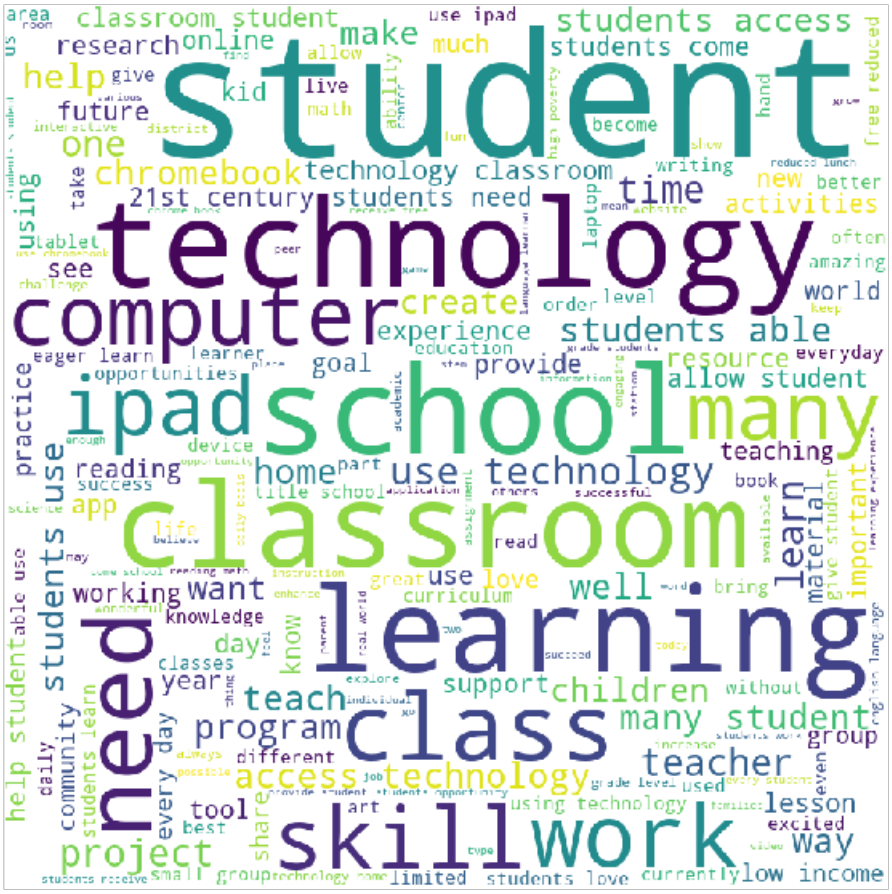

```
**********3 cluster**********
           id teacher_prefix school_state project_submitted_datetime  \
48084  p011894            Ms.           WI        2016-07-18 14:43:03
59356  p245562           Mrs.           MO        2017-02-07 18:51:07
93349  p236406            Ms.           ME        2016-09-19 22:01:59
26826  p185758            Ms.           KY        2016-11-14 19:57:32
84107  p214781           Mrs.           SC        2016-09-19 20:26:27

      project_grade_category project_subject_categories  \
48084            Grades_3_5        Literacy & Language
59356            Grades_6_8        Literacy & Language
93349            Grades_6_8        Literacy & Language
26826            Grades_3_5        Literacy & Language
84107         Grades_PreK_2        Literacy & Language

       project_subject_subcategories  \
48084                      Literacy
```

```
48084                         Literacy
59356                         Literacy
93349  Literacy, Literature & Writing
26826                         Literacy
84107                         Literacy

                             project_title  \
48084                  third grade book lovers
59356  engaging students families read together
93349              read your heart out with us
26826      student choice means student growth
84107                   the just right library

                            project_resource_summary  \
48084  my students need variety texts grow independen...
59356  my students need 7 fire tablets check books li...
93349  my students need hands books authors hopefully...
26826  my students need storage individualized differ...
84107  my students need leveled books support literac...

       teacher_number_of_previously_posted_projects  project_is_approved  \
48084                                             0                    1
59356                                            59                    1
93349                                             1                    1
26826                                             1                    0
84107                                             2                    1

                                                   essay   price  quantity
48084  i entering classroom wonderful unique third gr...  455.14         9
59356  do remember finding one special book could not...   39.99         7
93349  do love read heart we in february school host ...   20.83        21
26826  the students fifth grade classroom come divers...   64.48         4
84107  in kindergarten students come many different a...  461.43         3
```




```
********************4 cluster********************
          id teacher_prefix school_state project_submitted_datetime  \
1860   p115325           Mrs.           OR        2016-10-31 23:30:00
52041  p210239            Ms.           TX        2017-01-15 20:20:08
26116  p093362            Mr.           CA        2017-01-10 13:35:13
80292  p134644           Mrs.           LA        2016-12-11 15:40:18
72745  p226920           Mrs.           LA        2016-10-13 16:15:41

      project_grade_category          project_subject_categories  \
```

```
      project_grade_category           project_subject_categories  \
1860             Grades_3_5                     Music & The Arts
52041            Grades_3_5     History & Civics, Math & Science
26116            Grades_3_5  Math & Science, Literacy & Language
80292         Grades_PreK_2  Literacy & Language, Math & Science
72745         Grades_PreK_2                  Literacy & Language

              project_subject_subcategories  \
1860                             Visual Arts
52041        Financial Literacy, Mathematics
26116  Applied Sciences, Literature & Writing
80292                  Literacy, Mathematics
72745                   Literature & Writing

                                        project_title  \
1860                      clean surfaces watercolor art
52041  becoming financially literate with little help...
26116                                technology title 1
80292                                    techi firsties
72745                       classroom rug echo flooding

                             project_resource_summary  \
1860   my students need watercolor paper art cleaning...
52041  my students need tablets help construct busine...
26116  my students need chromebooks research create c...
80292  my students need 10 samsung galaxy tablets 10 ...
72745  my students need classroom rug extra seating a...

       teacher_number_of_previously_posted_projects  project_is_approved  \
1860                                              1                    1
52041                                             2                    1
26116                                             2                    1
80292                                             0                    1
72745                                             1                    1

                                                 essay   price  quantity
1860   my students love take pride rural community th...   50.07         8
52041  my students one thing common eagerness grow sc...  248.46         9
26116  our school one oldest district many students c...  175.42         5
80292  we young energetic group first graders come di...  650.92        30
72745  reading comprehension enjoyment books importan...  339.77         6
```

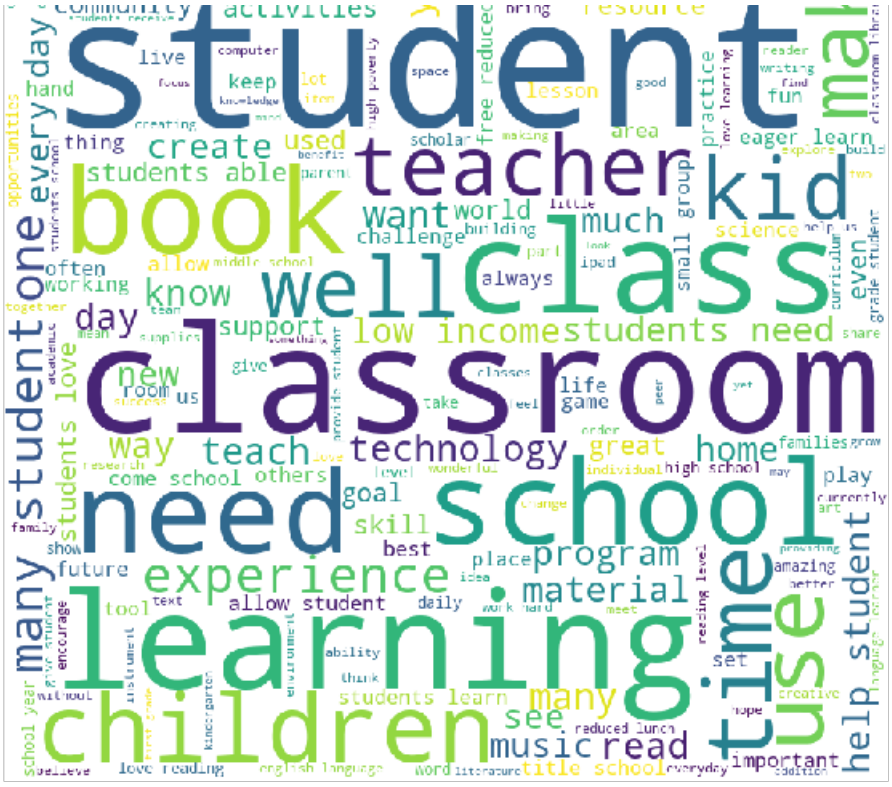


4. Applying AgglomerativeClustering

In [12]:

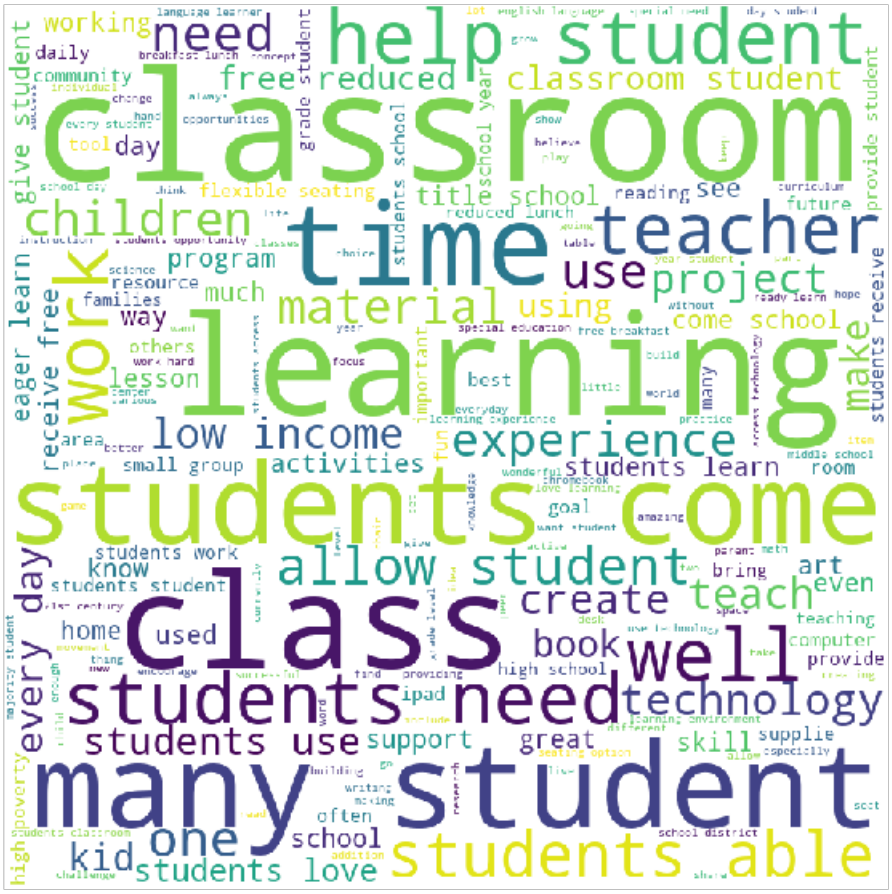
```
def infoClusters(clusters,predicted):
    for i in range(clusters):
        print(("*"*20)+str(i)+" cluster"+("*"*20))
        temp=(predicted==i)
        print(data[temp][:5])
        wordcloud(temp)


K = range(2,10,2)
for k in K:
    ac = AgglomerativeClustering(n_clusters=k)
    predicted = ac.fit_predict(x_train_new.toarray())
    print(("*"*20)+" Agglomerative clustring with clusters of "+str(k)+("*"*20))
    infoClusters(k,predicted)
```

```
******************** Agglomerative clustring with clusters of 2********************
********************0 cluster**********************
           id teacher_prefix school_state project_submitted_datetime  \
1860   p115325          Mrs.           OR        2016-10-31 23:30:00
35390  p158955           Ms.           CA        2016-07-22 16:42:42
48084  p011894           Ms.           WI        2016-07-18 14:43:03
52041  p210239           Ms.           TX        2017-01-15 20:20:08
59356  p245562          Mrs.           MO        2017-02-07 18:51:07

      project_grade_category        project_subject_categories  \
1860              Grades_3_5                Music & The Arts
35390             Grades_6_8                   Math & Science
48084             Grades_3_5               Literacy & Language
52041             Grades_3_5  History & Civics, Math & Science
59356             Grades_6_8               Literacy & Language

         project_subject_subcategories  \
1860                       Visual Arts
35390            Environmental Science
48084                         Literacy
52041  Financial Literacy, Mathematics
59356                         Literacy

                                      project_title  \
1860                      clean surfaces watercolor art
35390                                   click engage
48084                          third grade book lovers
52041  becoming financially literate with little help...
59356           engaging students families read together

                             project_resource_summary  \
1860   my students need watercolor paper art cleaning...
35390  my students need 34 interactive clickers make ...
48084  my students need variety texts grow independen...
52041  my students need tablets help construct busine...
59356  my students need 7 fire tablets check books li...

       teacher_number_of_previously_posted_projects  project_is_approved  \
1860                                              1                   1
35390                                             0                   1
48084                                             0                   1
52041                                             2                   1
59356                                            59                   1

                                                 essay    price  quantity
1860   my students love take pride rural community th...    50.07         8
35390  my middle school located underserved community...  1541.14         3
48084  i entering classroom wonderful unique third gr...   455.14         9
52041  my students one thing common eagerness grow sc...   248.46         9
59356  do remember finding one special book could not...    39.99         7
```



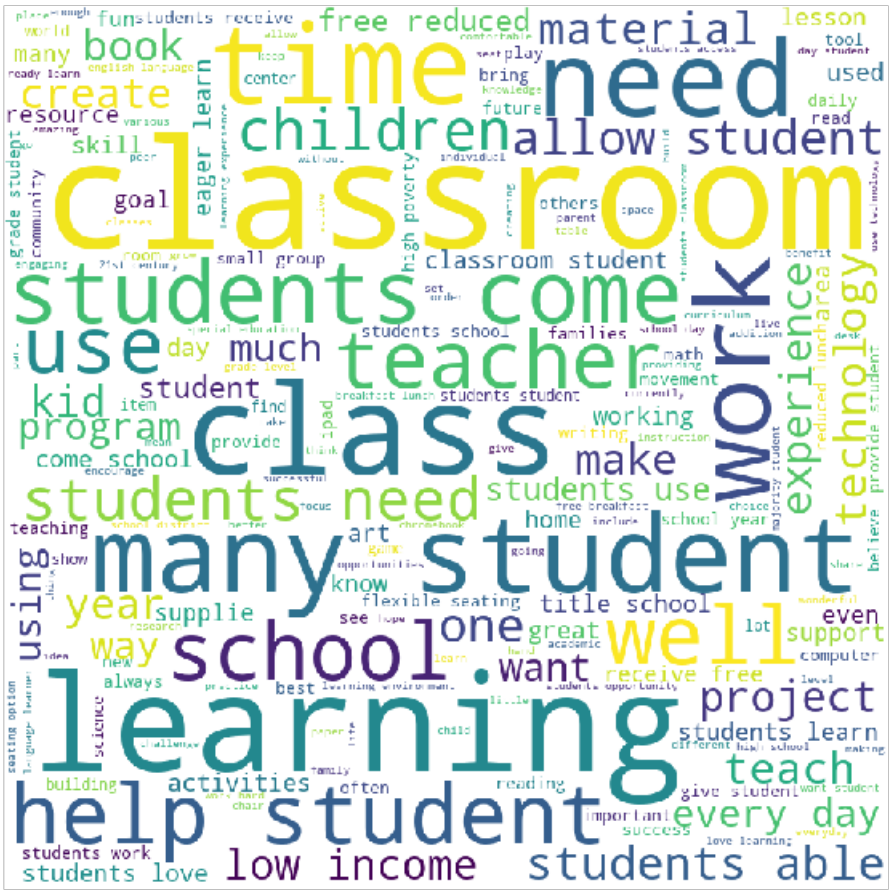

```
********************1 cluster********************
            id teacher_prefix school_state project_submitted_datetime  \
105447  p118457        Teacher           NC        2016-07-05 16:12:10
69262   p013133           Mrs.           NY        2016-10-15 16:53:46
55966   p205746            Ms.           UT        2016-08-25 01:29:34
88354   p047488            Ms.           HI        2016-07-01 17:49:36
55531   p042481           Mrs.           AL        2017-01-25 21:29:09

       project_grade_category     project_subject_categories  \
105447            Grades_9_12  Math & Science, Special Needs
69262              Grades_3_5                Special Needs
55966             Grades_9_12              History & Civics
88354              Grades_6_8                Math & Science
55531              Grades_3_5            Literacy & Language

                project_subject_subcategories  \
105447    Environmental Science, Special Needs
69262                          Special Needs
55966     History & Geography, Social Sciences
88354   Applied Sciences, Environmental Science
55531           Literacy, Literature & Writing

                                 project_title  \
105447                rain barrels school garden
69262                         wobbly learning
55966           hold tight using ipad case grip
88354   life science lab materials let investigate
55531         learning with the click of a mouse

                                project_resource_summary  \
105447  my students need rain barrels collect water ga...
69262   my students need five wobble chair they energy...
55966   my students need ipad case drop new ipad not b...
88354   my students need materials conduct biology lab...
55531   my students need benefit using usb wired mice ...

        teacher_number_of_previously_posted_projects  project_is_approved  \
105447                                             0                    1
69262                                             20                    0
55966                                              2                    1
88354                                              9                    1
55531                                              0                    1

                                                   essay   price  quantity
105447  environmental science students no idea food co...   93.69         4
69262   my students need six learniture active learnin...  209.85         5
55966   our high school school residing salt lake city...  102.31         2
```

```
55966   our high school school residing salt lake city...  102.31         2
88354   my students awesome they continually amaze fre...  188.53         8
55531   my school encompasses k 4th grade students we ...    6.29        26
```



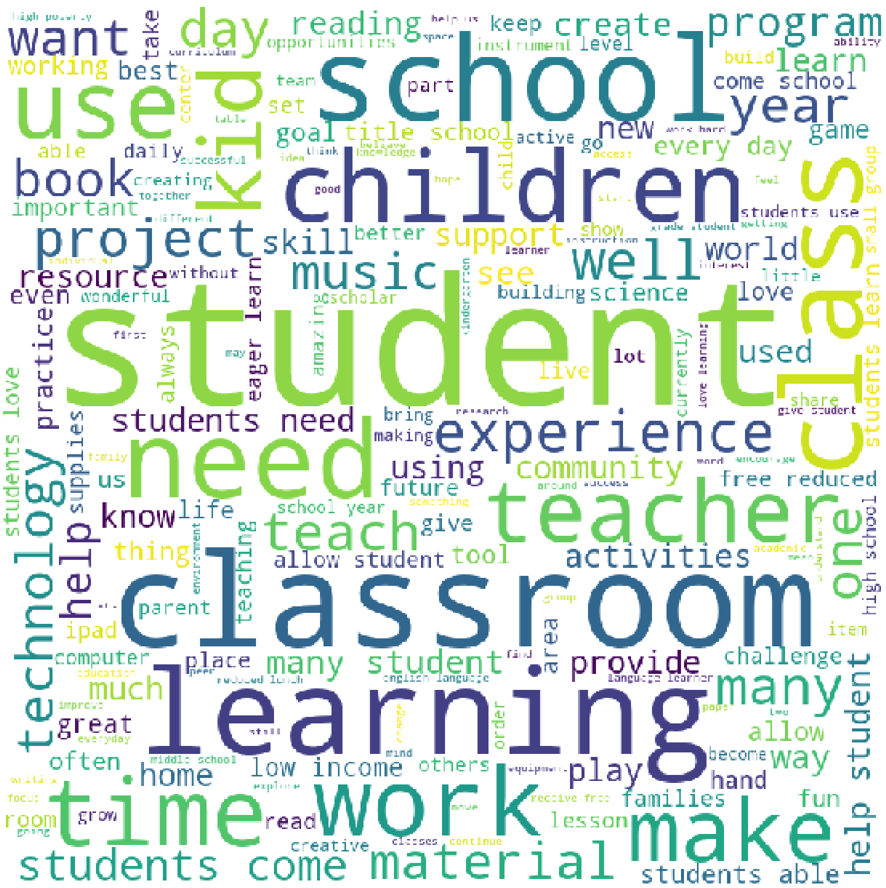

```
******************** Agglomerative clustring with clusters of 4**********************
**********************0 cluster**********************
            id teacher_prefix school_state project_submitted_datetime  \
105447  p118457        Teacher           NC        2016-07-05 16:12:10
69262   p013133           Mrs.           NY        2016-10-15 16:53:46
55966   p205746            Ms.           UT        2016-08-25 01:29:34
55531   p042481           Mrs.           AL        2017-01-25 21:29:09
102804  p036514           Mrs.           CA        2016-06-29 01:40:00

       project_grade_category          project_subject_categories  \
105447            Grades_9_12       Math & Science, Special Needs
69262              Grades_3_5                       Special Needs
55966             Grades_9_12                    History & Civics
55531              Grades_3_5                  Literacy & Language
102804          Grades_PreK_2  Literacy & Language, Math & Science

              project_subject_subcategories  \
105447  Environmental Science, Special Needs
69262                         Special Needs
55966   History & Geography, Social Sciences
55531         Literacy, Literature & Writing
102804                Literacy, Mathematics

                         project_title  \
105447          rain barrels school garden
69262                     wobbly learning
55966      hold tight using ipad case grip
55531   learning with the click of a mouse
102804    forget paper let go eco friendly

                              project_resource_summary  \
105447  my students need rain barrels collect water ga...
69262   my students need five wobble chair they energy...
55966   my students need ipad case drop new ipad not b...
55531   my students need benefit using usb wired mice ...
102804  my students need boogie boards inexpensive reu...

       teacher_number_of_previously_posted_projects  project_is_approved  \
105447                                            0                    1
```

```
69262                                                  20                    0
55966                                                   2                    1
55531                                                   0                    1
102804                                                 10                    1

                                                     essay   price  quantity
105447  environmental science students no idea food co...   93.69         4
69262   my students need six learniture active learnin...  209.85         5
55966   our high school school residing salt lake city...  102.31         2
55531   my school encompasses k 4th grade students we ...    6.29        26
102804  i classroom wonderful students we start learni...   29.99         4
```



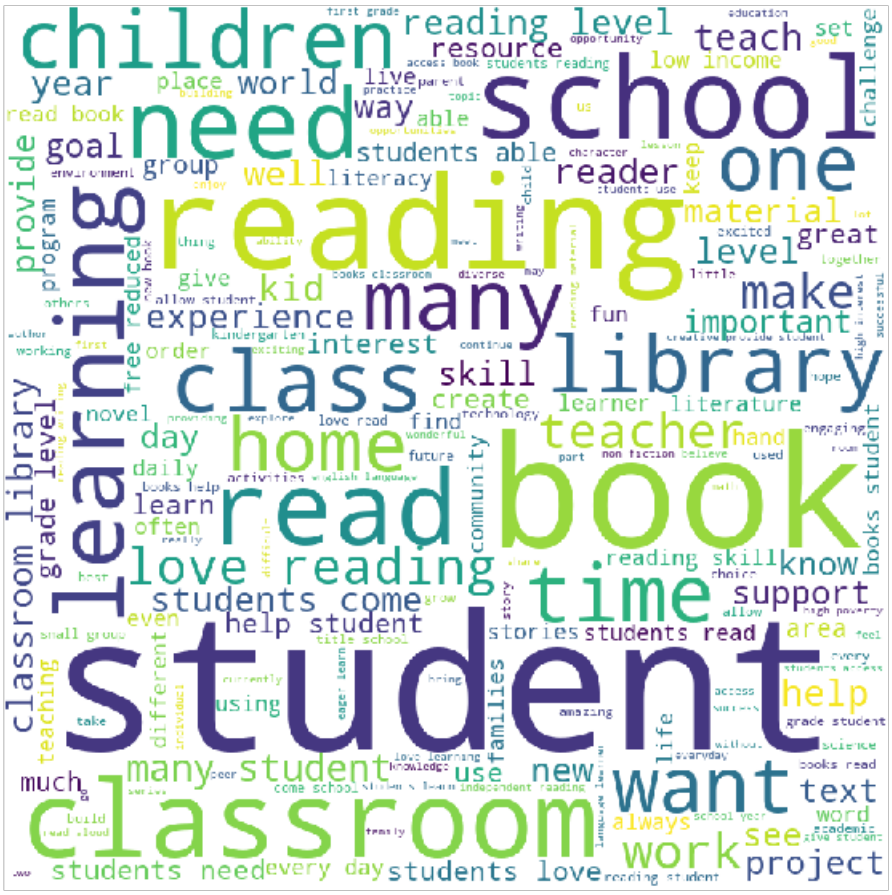

```
********************1 cluster********************
            id teacher_prefix school_state project_submitted_datetime  \
1860   p115325          Mrs.           OR        2016-10-31 23:30:00
35390  p158955           Ms.           CA        2016-07-22 16:42:42
52041  p210239           Ms.           TX        2017-01-15 20:20:08
24971  p254758           Mr.           MD        2016-06-20 10:14:16
22001  p051016          Mrs.           CO        2016-09-13 16:50:42

      project_grade_category           project_subject_categories  \
1860              Grades_3_5                     Music & The Arts
35390             Grades_6_8                      Math & Science
52041             Grades_3_5     History & Civics, Math & Science
24971             Grades_3_5                      Math & Science
22001             Grades_3_5  Literacy & Language, Math & Science

           project_subject_subcategories  \
1860                         Visual Arts
35390              Environmental Science
52041    Financial Literacy, Mathematics
24971              Environmental Science
22001  Literature & Writing, Mathematics

                                        project_title  \
1860                      clean surfaces watercolor art
35390                                      click engage
52041  becoming financially literate with little help...
24971                                 how plants help us
22001                                engaging activities

                              project_resource_summary  \
```

```
1860   my students need watercolor paper art cleaning...
35390  my students need 34 interactive clickers make ...
52041  my students need tablets help construct busine...
24971  my students need lab equipment test plants imp...
22001  my students need multiplication division bingo...

       teacher_number_of_previously_posted_projects  project_is_approved  \
1860                                              1                    1
35390                                             0                    1
52041                                             2                    1
24971                                            12                    1
22001                                             2                    1

                                                   essay    price  quantity
1860   my students love take pride rural community th...    50.07         8
35390  my middle school located underserved community...  1541.14         3
52041  my students one thing common eagerness grow sc...   248.46         9
24971  we fabulous group fifth grade students baltimo...   389.64         3
22001  this year i opportunity move kindergarten 5th ...   174.94         6
```



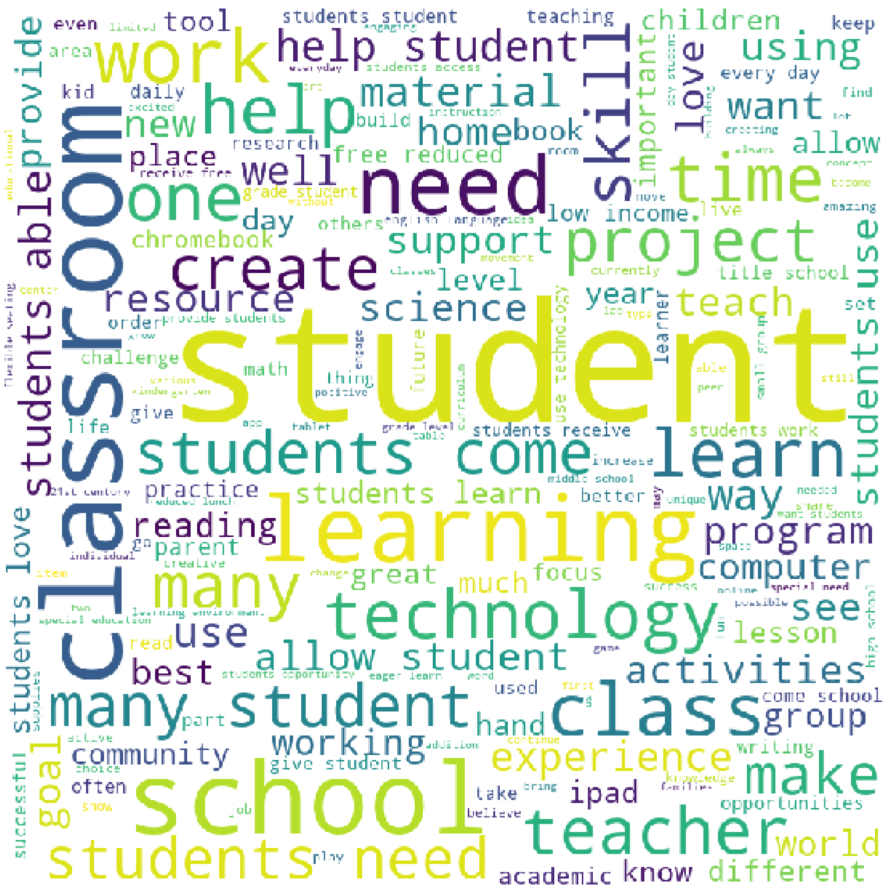

```
*******************2 cluster*******************
           id teacher_prefix school_state project_submitted_datetime  \
48084  p011894            Ms.           WI        2016-07-18 14:43:03
59356  p245562           Mrs.           MO        2017-02-07 18:51:07
59452  p226149           Mrs.           ME        2017-03-07 13:01:09
93349  p236406            Ms.           ME        2016-09-19 22:01:59
26826  p185758            Ms.           KY        2016-11-14 19:57:32

      project_grade_category project_subject_categories  \
48084            Grades_3_5         Literacy & Language
59356            Grades_6_8         Literacy & Language
59452         Grades_PreK_2         Literacy & Language
93349            Grades_6_8         Literacy & Language
26826            Grades_3_5         Literacy & Language

        project_subject_subcategories  \
48084                        Literacy
59356                        Literacy
59452                        Literacy
93349  Literacy, Literature & Writing
26826                        Literacy
```

```
                                project_title  \
48084                   third grade book lovers
59356  engaging students families read together
59452                   spruce seuss collection
93349                read your heart out with us
26826       student choice means student growth

                               project_resource_summary  \
48084  my students need variety texts grow independen...
59356  my students need 7 fire tablets check books li...
59452  my students need dr seuss books add learning l...
93349  my students need hands books authors hopefully...
26826  my students need storage individualized differ...

       teacher_number_of_previously_posted_projects  project_is_approved  \
48084                                             0                    1
59356                                            59                    1
59452                                            21                    1
93349                                             1                    1
26826                                             1                    0

                                                   essay   price  quantity
48084  i entering classroom wonderful unique third gr...  455.14         9
59356  do remember finding one special book could not...   39.99         7
59452  my students come various backgrounds we teach ...  192.94        28
93349  do love read heart we in february school host ...   20.83        21
26826  the students fifth grade classroom come divers...   64.48         4
```



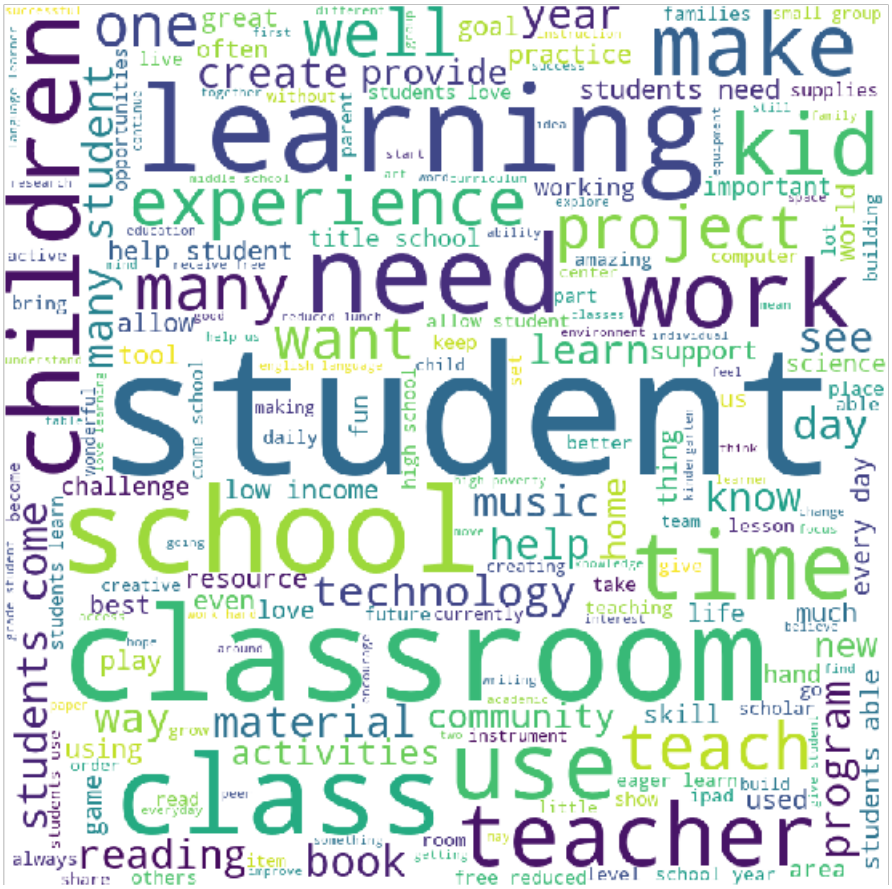

```
**********3 cluster**********
            id teacher_prefix school_state project_submitted_datetime  \
88354   p047488           Ms.           HI        2016-07-01 17:49:36
42921   p184874       Teacher           IL        2016-10-26 21:13:32
83483   p153579           Mr.           SC        2016-08-12 08:17:47
36330   p259741          Mrs.           TX        2016-09-25 22:18:38
106956  p054463           Mr.           NY        2016-10-19 16:16:09

       project_grade_category            project_subject_categories  \
88354             Grades_6_8                         Math & Science
42921             Grades_6_8  Math & Science, Literacy & Language
83483             Grades_3_5  Literacy & Language, Music & The Arts
36330          Grades_PreK_2  Literacy & Language, Math & Science
106956          Grades_9_12        Math & Science, Special Needs
```

```
                       project_subject_subcategories  \
88354         Applied Sciences, Environmental Science
42921   Health & Life Science, Literature & Writing
83483                                Literacy, Music
36330                          Literacy, Mathematics
106956                     Mathematics, Special Needs

                                          project_title  \
88354         life science lab materials let investigate
42921           flipping our classroom into 21st century
83483                                 minds with mallets
36330   technology is giving students chance dream big
106956            individualize differentiate technology

                                 project_resource_summary  \
88354   my students need materials conduct biology lab...
42921   my students need chromebooks flip classroom 21...
83483   my students need increase number orff instrume...
36330   my students need laptops things make successfu...
106956  my students need two chromebooks individualize...

        teacher_number_of_previously_posted_projects  project_is_approved  \
88354                                              9                    1
42921                                              3                    1
83483                                              1                    0
36330                                              0                    1
106956                                             7                    1

                                                    essay   price  quantity
88354   my students awesome they continually amaze fre...  188.53         8
42921   as teacher low income high poverty school dist...  179.00         4
83483   i teach music 600 students school i love worki...  955.50         4
36330   i teach title i elementary school most student...  199.99         4
106956  my school community tight knit students bond m...  159.43         2
```

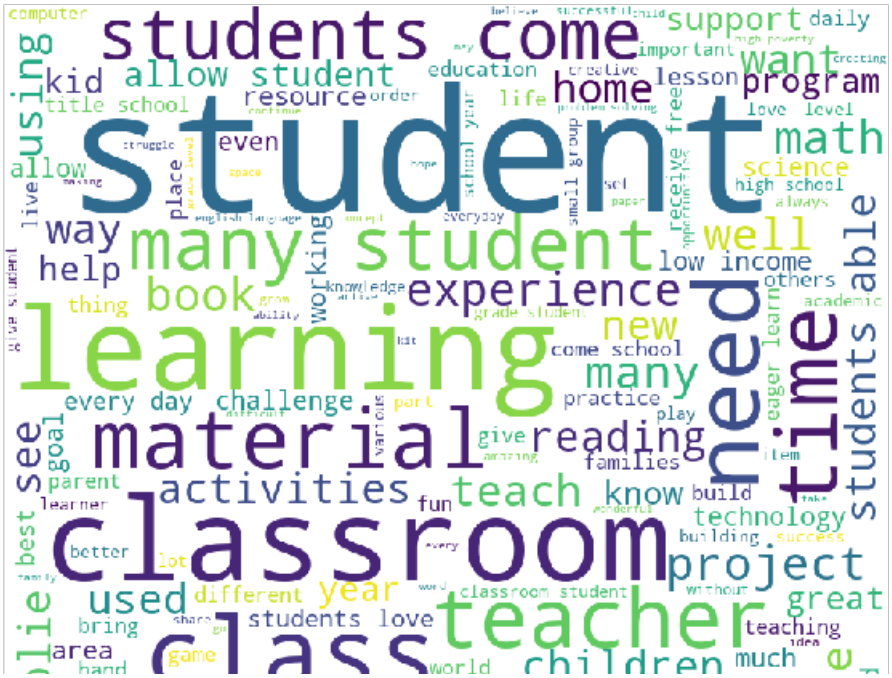


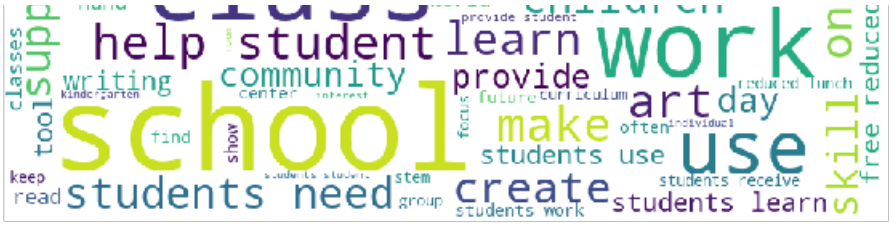
```
******************* Agglomerative clustring with clusters of 6*******************
*******************0 cluster*******************
           id teacher_prefix school_state project_submitted_datetime  \
1860   p115325           Mrs.           OR        2016-10-31 23:30:00
35390  p158955            Ms.           CA        2016-07-22 16:42:42
52041  p210239            Ms.           TX        2017-01-15 20:20:08
24971  p254758            Mr.           MD        2016-06-20 10:14:16
```

```
22001  p051016           Mrs.           CO        2016-09-13 16:50:42

      project_grade_category           project_subject_categories  \
1860              Grades_3_5                      Music & The Arts
35390             Grades_6_8                        Math & Science
52041             Grades_3_5     History & Civics, Math & Science
24971             Grades_3_5                        Math & Science
22001             Grades_3_5  Literature & Language, Math & Science

          project_subject_subcategories  \
1860                         Visual Arts
35390              Environmental Science
52041    Financial Literacy, Mathematics
24971              Environmental Science
22001  Literature & Writing, Mathematics

                                      project_title  \
1860                      clean surfaces watercolor art
35390                                    click engage
52041  becoming financially literate with little help...
24971                                how plants help us
22001                              engaging activities

                              project_resource_summary  \
1860   my students need watercolor paper art cleaning...
35390  my students need 34 interactive clickers make ...
52041  my students need tablets help construct busine...
24971  my students need lab equipment test plants imp...
22001  my students need multiplication division bingo...

       teacher_number_of_previously_posted_projects  project_is_approved  \
1860                                              1                    1
35390                                             0                    1
52041                                             2                    1
24971                                            12                    1
22001                                             2                    1

                                                   essay    price  quantity
1860   my students love take pride rural community th...    50.07         8
35390  my middle school located underserved community...  1541.14         3
52041  my students one thing common eagerness grow sc...   248.46         9
24971  we fabulous group fifth grade students baltimo...   389.64         3
22001  this year i opportunity move kindergarten 5th ...   174.94         6
```



```
********************1 cluster********************
            id teacher_prefix school_state project_submitted_datetime  \
105447  p118457        Teacher           NC        2016-07-05 16:12:10
102804  p036514           Mrs.           CA        2016-06-29 01:40:00
67213   p102993           Mrs.           NV        2017-02-05 17:18:18
100478  p140242            Ms.           CT        2016-09-19 18:04:30
106451  p096584           Mrs.           CA        2016-08-10 01:55:26

       project_grade_category           project_subject_categories  \
105447           Grades_9_12        Math & Science, Special Needs
102804          Grades_PreK_2  Literacy & Language, Math & Science
67213           Grades_PreK_2   Literacy & Language, Special Needs
100478          Grades_PreK_2   Literacy & Language, Special Needs
106451             Grades_3_5                         Math & Science

              project_subject_subcategories  \
105447  Environmental Science, Special Needs
102804                 Literacy, Mathematics
67213                Literacy, Special Needs
100478               Literacy, Special Needs
106451         Applied Sciences, Mathematics

                                        project_title  \
105447                       rain barrels school garden
102804                 forget paper let go eco friendly
67213                 environment engagement excitement
100478  please help my students with autism practice t...
106451            engaging students with robotics coding

                                project_resource_summary  \
105447  my students need rain barrels collect water ga...
102804  my students need boogie boards inexpensive reu...
67213   my students need florescent light filters acad...
100478  my students need learning tools reinforce read...
106451  my students need ozobot robots help expand ste...

        teacher_number_of_previously_posted_projects  project_is_approved  \
105447                                             0                    1
102804                                            10                    1
67213                                              4                    1
100478                                             2                    0
106451                                             1                    1

                                                   essay   price  quantity
105447  environmental science students no idea food co...   93.69         4
102804  i classroom wonderful students we start learni...   29.99         4
67213   my students energetic kind curious children th...  234.16         7
100478  i teach classroom 8 students autism a typical ...  269.95         6
106451  i teach students 10 12 years old eighty percen...  119.99         4
```





**********2 cluster**********

	id	teacher_prefix	school_state	project_submitted_datetime
48084	p011894	Ms.	WI	2016-07-18 14:43:03
59356	p245562	Mrs.	MO	2017-02-07 18:51:07
59452	p226149	Mrs.	ME	2017-03-07 13:01:09
93349	p236406	Ms.	ME	2016-09-19 22:01:59
26826	p185758	Ms.	KY	2016-11-14 19:57:32

	project_grade_category	project_subject_categories
48084	Grades_3_5	Literacy & Language
59356	Grades_6_8	Literacy & Language
59452	Grades_PreK_2	Literacy & Language
93349	Grades_6_8	Literacy & Language
26826	Grades_3_5	Literacy & Language

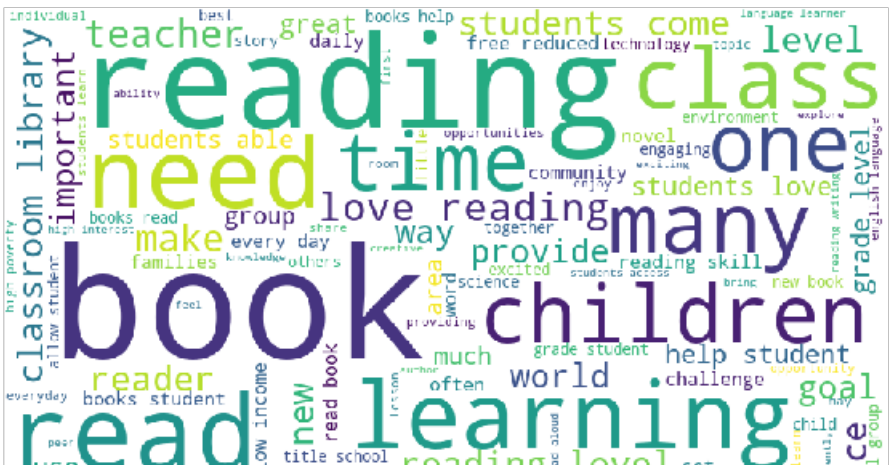
	project_subject_subcategories
48084	Literacy
59356	Literacy
59452	Literacy
93349	Literacy, Literature & Writing
26826	Literacy

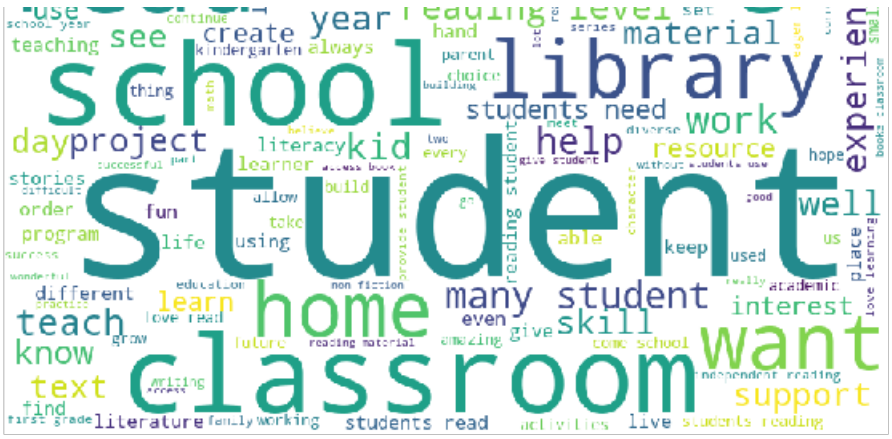
	project_title
48084	third grade book lovers
59356	engaging students families read together
59452	spruce seuss collection
93349	read your heart out with us
26826	student choice means student growth

	project_resource_summary
48084	my students need variety texts grow independen...
59356	my students need 7 fire tablets check books li...
59452	my students need dr seuss books add learning l...
93349	my students need hands books authors hopefully...
26826	my students need storage individualized differ...

	teacher_number_of_previously_posted_projects	project_is_approved
48084	0	1
59356	59	1
59452	21	1
93349	1	1
26826	1	0

	essay	price	quantity
48084	i entering classroom wonderful unique third gr...	455.14	9
59356	do remember finding one special book could not...	39.99	7
59452	my students come various backgrounds we teach ...	192.94	28
93349	do love read heart we in february school host ...	20.83	21
26826	the students fifth grade classroom come divers...	64.48	4





**********3 cluster**********

	id	teacher_prefix	school_state	project_submitted_datetime	\
88354	p047488	Ms.	HI	2016-07-01 17:49:36	
42921	p184874	Teacher	IL	2016-10-26 21:13:32	
83483	p153579	Mr.	SC	2016-08-12 08:17:47	
36330	p259741	Mrs.	TX	2016-09-25 22:18:38	
106956	p054463	Mr.	NY	2016-10-19 16:16:09	

	project_grade_category	project_subject_categories	\
88354	Grades_6_8	Math & Science	
42921	Grades_6_8	Math & Science, Literacy & Language	
83483	Grades_3_5	Literacy & Language, Music & The Arts	
36330	Grades_PreK_2	Literacy & Language, Math & Science	
106956	Grades_9_12	Math & Science, Special Needs	

	project_subject_subcategories	\
88354	Applied Sciences, Environmental Science	
42921	Health & Life Science, Literature & Writing	
83483	Literacy, Music	
36330	Literacy, Mathematics	
106956	Mathematics, Special Needs	

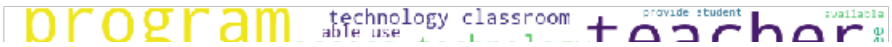
	project_title	\
88354	life science lab materials let investigate	
42921	flipping our classroom into 21st century	
83483	minds with mallets	
36330	technology is giving students chance dream big	
106956	individualize differentiate technology	

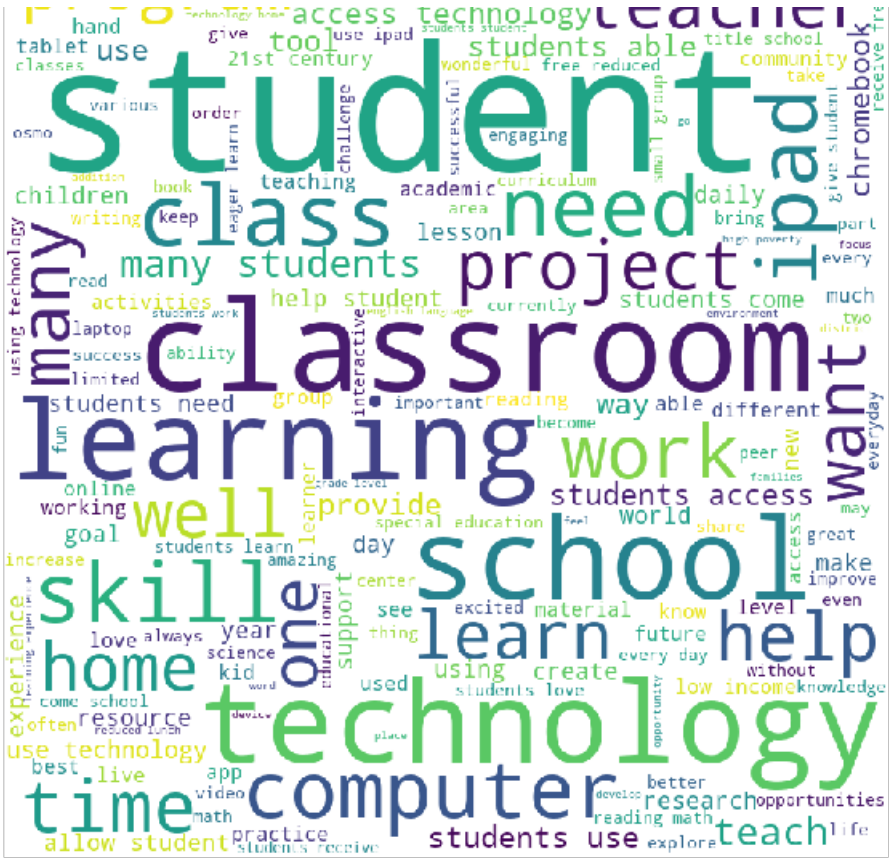
	project_resource_summary	\
88354	my students need materials conduct biology lab...	
42921	my students need chromebooks flip classroom 21...	
83483	my students need increase number orff instrume...	
36330	my students need laptops things make successfu...	
106956	my students need two chromebooks individualize...	

	teacher_number_of_previously_posted_projects	project_is_approved	\
88354	9	1	
42921	3	1	
83483	1	0	
36330	0	1	
106956	7	1	

	essay	price	quantity
88354	my students awesome they continually amaze fre...	188.53	8
42921	as teacher low income high poverty school dist...	179.00	4
83483	i teach music 600 students school i love worki...	955.50	4
36330	i teach title i elementary school most student...	199.99	4
106956	my school community tight knit students bond m...	159.43	2

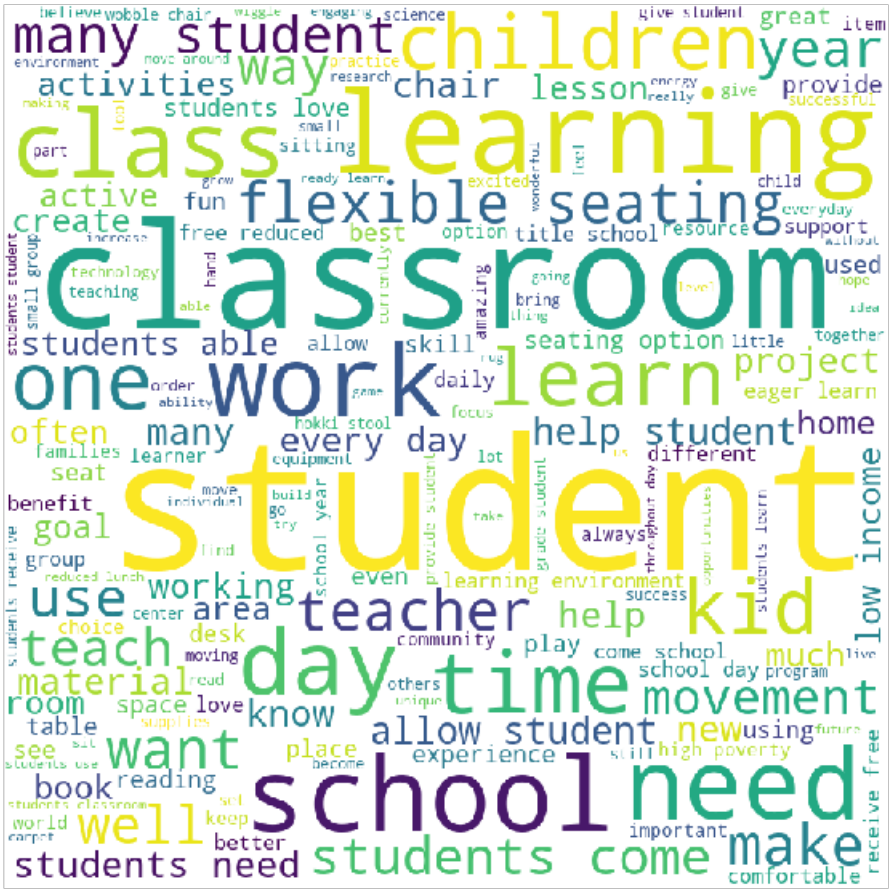
```
**********4 cluster**********
            id teacher_prefix school_state project_submitted_datetime  \
55966   p205746            Ms.           UT        2016-08-25 01:29:34
55531   p042481           Mrs.           AL        2017-01-25 21:29:09
26116   p093362            Mr.           CA        2017-01-10 13:35:13
46770   p026359           Mrs.           OK        2016-09-14 16:01:00
80292   p134644           Mrs.           LA        2016-12-11 15:40:18

      project_grade_category           project_subject_categories  \
55966            Grades_9_12                    History & Civics
55531             Grades_3_5                Literacy & Language
26116             Grades_3_5  Math & Science, Literacy & Language
46770             Grades_3_5      Applied Learning, Special Needs
80292          Grades_PreK_2  Literacy & Language, Math & Science

              project_subject_subcategories  \
55966    History & Geography, Social Sciences
55531          Literacy, Literature & Writing
26116  Applied Sciences, Literature & Writing
46770        Early Development, Special Needs
80292                    Literacy, Mathematics

                             project_title  \
55966         hold tight using ipad case grip
55531      learning with the click of a mouse
26116                    technology title 1
46770  speech language special needs children
80292                         techi firsties

                             project_resource_summary  \
55966  my students need ipad case drop new ipad not b...
55531  my students need benefit using usb wired mice ...
26116  my students need chromebooks research create c...
46770  my students need ipad air 2 access speech lang...
80292  my students need 10 samsung galaxy tablets 10 ...

       teacher_number_of_previously_posted_projects  project_is_approved  \
55966                                             2                   1
55531                                             0                   1
26116                                             2                   1
46770                                             0                   1
80292                                             0                   1

                                                   essay   price  quantity
55966  our high school school residing salt lake city...  102.31         2
55531  my school encompasses k 4th grade students we ...    6.29        26
26116  our school one oldest district many students c...  175.42         5
46770  i serve students special needs the area live h...  551.23         2
80292  we young energetic group first graders come di...  650.92        30
```





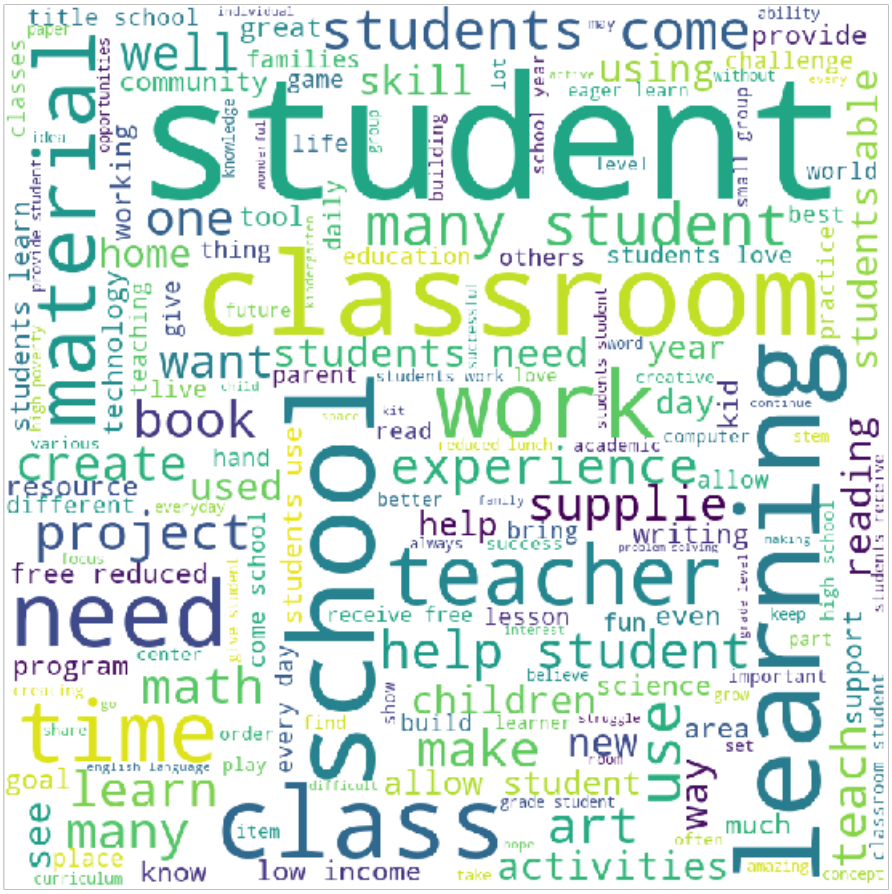

```
**********5 cluster**********
           id teacher_prefix school_state project_submitted_datetime  \
69262  p013133           Mrs.           NY        2016-10-15 16:53:46
59836  p056013           Mrs.           TX        2017-01-03 22:43:46
10776  p062274            Ms.           KY        2017-03-28 18:22:57
10240  p008999            Mr.           MA        2016-08-14 00:00:23
274    p006577            Ms.           IL        2016-07-25 17:14:02

      project_grade_category            project_subject_categories  \
69262            Grades_3_5                       Special Needs
59836         Grades_PreK_2  Applied Learning, Literacy & Language
10776         Grades_PreK_2    Literacy & Language, Math & Science
10240            Grades_3_5                        Health & Sports
274            Grades_9_12                         Math & Science

                   project_subject_subcategories  \
69262                                Special Needs
59836                  Early Development, Literacy
10776                        Literacy, Mathematics
10240            Gym & Fitness, Health & Wellness
274    Environmental Science, Health & Life Science

                                      project_title  \
69262                                wobbly learning
59836                       wiggle wiggle wiggle
10776                           wobbling working
10240                warming up with fitness gaming
274    classroom pets fish tadpoles turtles chameleon...

                           project_resource_summary  \
69262  my students need five wobble chair they energy...
59836  my students need hokki stools help get wiggles...
10776  my students need 5 wobble stools replace yoga ...
10240  my students need playstation 4 playstation cam...
274    my students need goldfish tropical fish tadpol...

       teacher_number_of_previously_posted_projects  project_is_approved  \
69262                                            20                   0
59836                                            18                   1
10776                                             5                   1
10240                                            74                   1
274                                              49                   1

                                                 essay   price  quantity
```

```
69262  my students need six learniture active learnin...  209.85          5
59836  i blessed teach fun loving group five year old...  392.03          9
10776  our classroom helps get wiggles i work this co...   50.12          5
10240  what learning today that something i hear stud...  672.03         13
274    my students freshmen co taught biology classro...  425.16         49
```



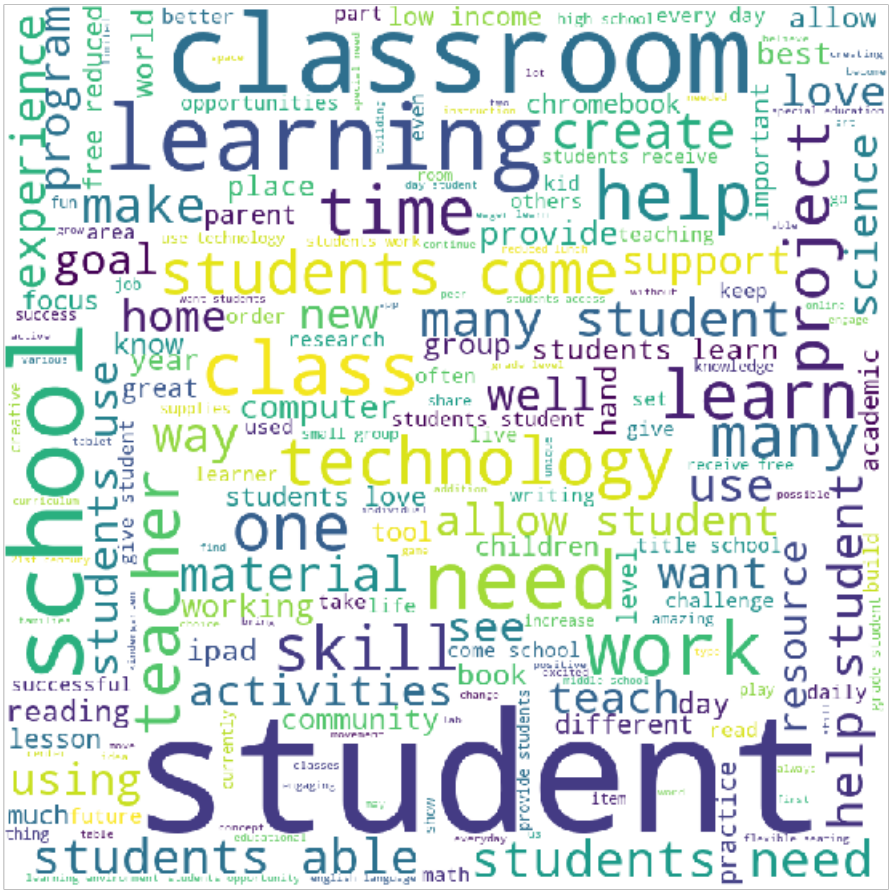

```
******************** Agglomerative clustring with clusters of 8**********************
*********************0 cluster**********************
             id teacher_prefix school_state project_submitted_datetime  \
105447  p118457        Teacher           NC        2016-07-05 16:12:10
102804  p036514           Mrs.           CA        2016-06-29 01:40:00
67213   p102993           Mrs.           NV        2017-02-05 17:18:18
100478  p140242            Ms.           CT        2016-09-19 18:04:30
106451  p096584           Mrs.           CA        2016-08-10 01:55:26

       project_grade_category           project_subject_categories  \
105447            Grades_9_12        Math & Science, Special Needs
102804          Grades_PreK_2  Literacy & Language, Math & Science
67213           Grades_PreK_2   Literacy & Language, Special Needs
100478          Grades_PreK_2   Literacy & Language, Special Needs
106451             Grades_3_5                        Math & Science

              project_subject_subcategories  \
105447  Environmental Science, Special Needs
102804                 Literacy, Mathematics
67213                Literacy, Special Needs
100478               Literacy, Special Needs
106451         Applied Sciences, Mathematics

                                       project_title  \
105447                      rain barrels school garden
102804                forget paper let go eco friendly
67213                environment engagement excitement
100478  please help my students with autism practice t...
106451            engaging students with robotics coding

                                project_resource_summary  \
105447  my students need rain barrels collect water ga...
102804  my students need boogie boards inexpensive reu...
67213   my students need florescent light filters acad...
100478  my students need learning tools reinforce read...
106451  my students need ozobot robots help expand ste...
```

```
        teacher_number_of_previously_posted_projects  project_is_approved  \
105447                                             0                    1
102804                                            10                    1
67213                                              4                    1
100478                                             2                    0
106451                                             1                    1

                                                    essay   price  quantity
105447  environmental science students no idea food co...   93.69         4
102804  i classroom wonderful students we start learni...   29.99         4
67213   my students energetic kind curious children th...  234.16         7
100478  i teach classroom 8 students autism a typical ...  269.95         6
106451  i teach students 10 12 years old eighty percen...  119.99         4
```



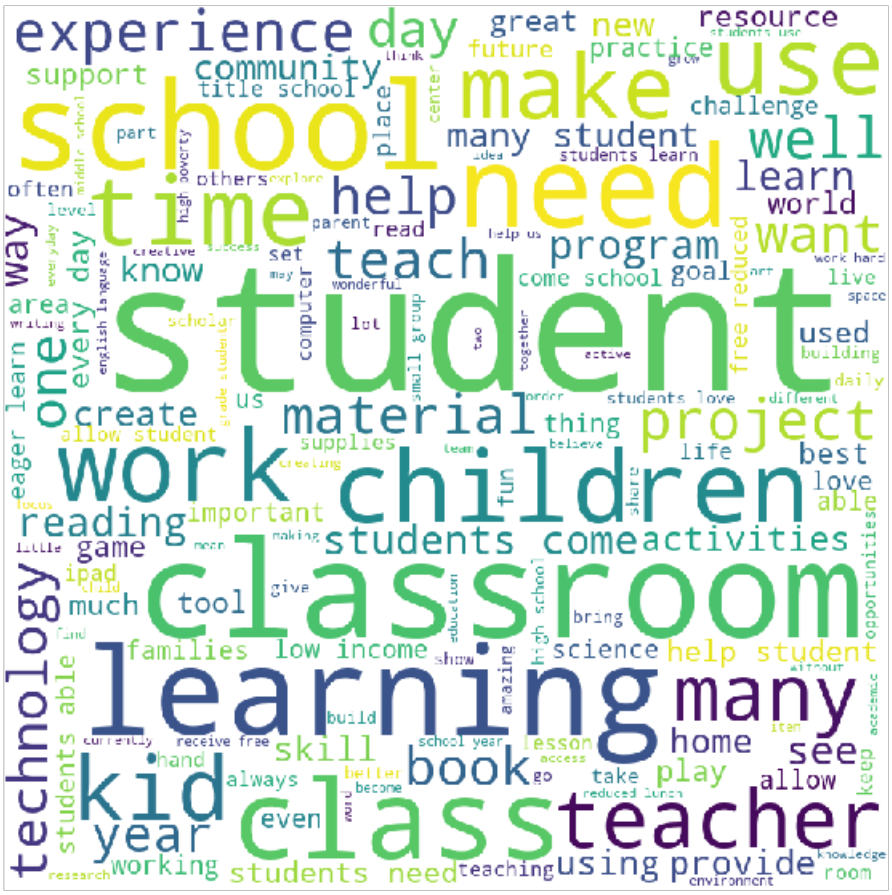

```
********************1 cluster*********************
            id teacher_prefix school_state project_submitted_datetime  \
88354   p047488            Ms.           HI        2016-07-01 17:49:36
42921   p184874        Teacher           IL        2016-10-26 21:13:32
83483   p153579            Mr.           SC        2016-08-12 08:17:47
36330   p259741           Mrs.           TX        2016-09-25 22:18:38
106956  p054463            Mr.           NY        2016-10-19 16:16:09

       project_grade_category            project_subject_categories  \
88354             Grades_6_8                       Math & Science
42921             Grades_6_8  Math & Science, Literacy & Language
83483             Grades_3_5  Literacy & Language, Music & The Arts
36330          Grades_PreK_2  Literacy & Language, Math & Science
106956           Grades_9_12        Math & Science, Special Needs

                    project_subject_subcategories  \
88354       Applied Sciences, Environmental Science
42921   Health & Life Science, Literature & Writing
83483                             Literacy, Music
36330                         Literacy, Mathematics
106956                   Mathematics, Special Needs

                                      project_title  \
88354       life science lab materials let investigate
42921         flipping our classroom into 21st century
83483                               minds with mallets
36330   technology is giving students chance dream big
```

```
106956          individualize differentiate technology

                                project_resource_summary  \
88354   my students need materials conduct biology lab...
42921   my students need chromebooks flip classroom 21...
83483   my students need increase number orff instrume...
36330   my students need laptops things make successfu...
106956  my students need two chromebooks individualize...

        teacher_number_of_previously_posted_projects  project_is_approved  \
88354                                              9                    1
42921                                              3                    1
83483                                              1                    0
36330                                              0                    1
106956                                             7                    1

                                                    essay   price  quantity
88354   my students awesome they continually amaze fre...  188.53         8
42921   as teacher low income high poverty school dist...  179.00         4
83483   i teach music 600 students school i love worki...  955.50         4
36330   i teach title i elementary school most student...  199.99         4
106956  my school community tight knit students bond m...  159.43         2
```



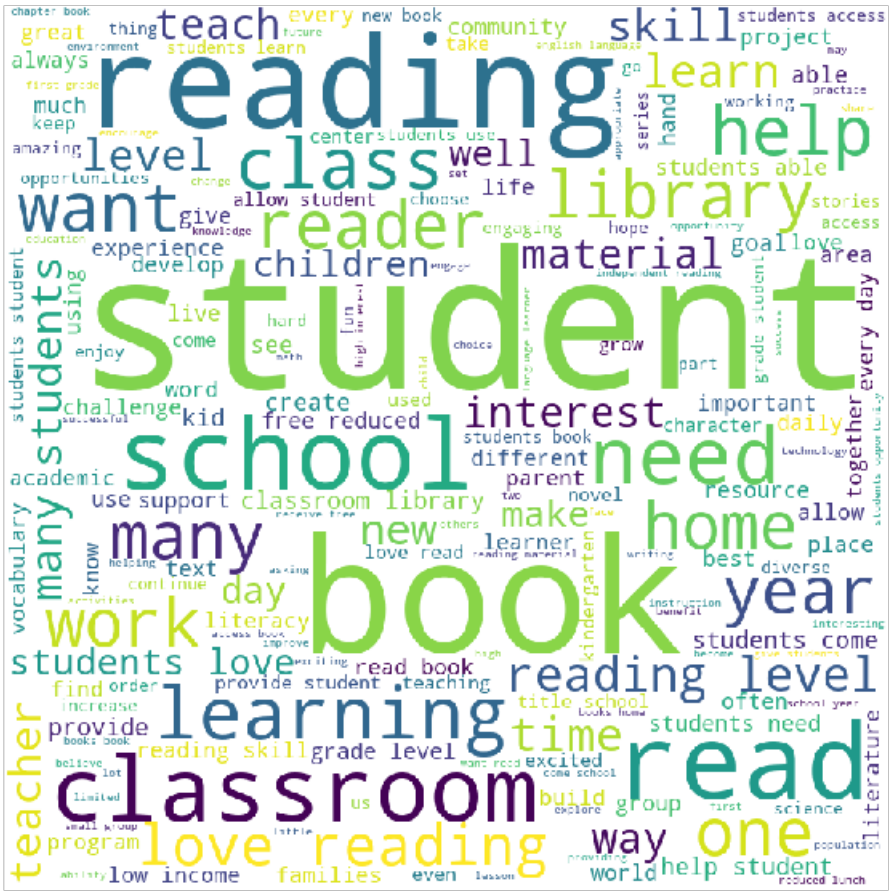

```
**********2 cluster**********
           id teacher_prefix school_state project_submitted_datetime  \
1860   p115325          Mrs.           OR        2016-10-31 23:30:00
35390  p158955           Ms.           CA        2016-07-22 16:42:42
52041  p210239           Ms.           TX        2017-01-15 20:20:08
24971  p254758           Mr.           MD        2016-06-20 10:14:16
22001  p051016          Mrs.           CO        2016-09-13 16:50:42

      project_grade_category          project_subject_categories  \
1860              Grades_3_5                       Music & The Arts
35390             Grades_6_8                         Math & Science
52041             Grades_3_5     History & Civics, Math & Science
24971             Grades_3_5                         Math & Science
22001             Grades_3_5  Literacy & Language, Math & Science

          project_subject_subcategories  \
1860                          Visual Arts
35390              Environmental Science
52041    Financial Literacy, Mathematics
```

```
24971              Environmental Science
22001  Literature & Writing, Mathematics

                                       project_title  \
1860                     clean surfaces watercolor art
35390                                    click engage
52041  becoming financially literate with little help...
24971                                how plants help us
22001                              engaging activities

                              project_resource_summary  \
1860   my students need watercolor paper art cleaning...
35390  my students need 34 interactive clickers make ...
52041  my students need tablets help construct busine...
24971  my students need lab equipment test plants imp...
22001  my students need multiplication division bingo...

       teacher_number_of_previously_posted_projects  project_is_approved  \
1860                                              1                    1
35390                                             0                    1
52041                                             2                    1
24971                                            12                    1
22001                                             2                    1

                                                   essay    price  quantity
1860   my students love take pride rural community th...    50.07         8
35390  my middle school located underserved community...  1541.14         3
52041  my students one thing common eagerness grow sc...   248.46         9
24971  we fabulous group fifth grade students baltimo...   389.64         3
22001  this year i opportunity move kindergarten 5th ...   174.94         6
```



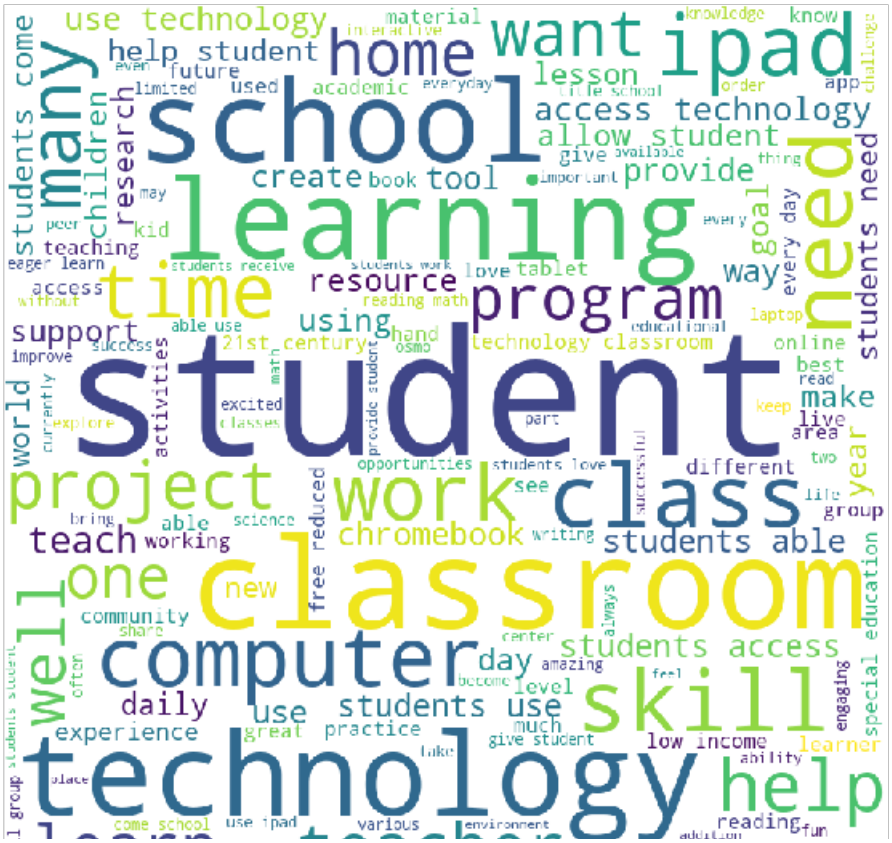

```
********************3 cluster********************
            id teacher_prefix school_state project_submitted_datetime  \
59356   p245562           Mrs.           MO        2017-02-07 18:51:07
60247   p035807           Mrs.           KS        2017-01-11 15:00:45
15985   p171338           Mrs.           CA        2017-02-13 00:20:44
9274    p033024            Ms.           FL        2016-10-27 17:17:42
83180   p197613           Mrs.           MD        2017-03-14 17:23:55

      project_grade_category          project_subject_categories  \
59356             Grades_6_8                Literacy & Language
60247          Grades_PreK_2                Literacy & Language
15985             Grades_3_5  Literacy & Language, Special Needs
```

```
9274              Grades_6_8                 Literacy & Language
83180          Grades_PreK_2                 Literacy & Language

            project_subject_subcategories  \
59356                             Literacy
60247                 Literature & Writing
15985  Literature & Writing, Special Needs
9274                              Literacy
83180                             Literacy

                                        project_title  \
59356           engaging students families read together
60247                         help foster love reading
15985                                   going chrome
9274   creating life long book lovers through best se...
83180     help us improve literacy skills engaging books

                                project_resource_summary  \
59356  my students need 7 fire tablets check books li...
60247  my students need need diverse books foster lov...
15985  my students need two acer chrome books develop...
9274   my students need variety book options interest...
83180  my students need engaging books level help lea...

       teacher_number_of_previously_posted_projects  project_is_approved  \
59356                                            59                    1
60247                                             0                    1
15985                                             2                    1
9274                                              1                    1
83180                                             7                    1

                                                   essay   price  quantity
59356  do remember finding one special book could not...   39.99         7
60247  my students part 2nd grade classroom high pove...  101.80        68
15985  i thirteen highly motivated special education ...  277.42         2
9274   at front room celebration bulletin board i pos...  192.68        27
83180  i currently 4th grade teacher switching 1st gr...  187.09        50
```




```
********************4 cluster********************
          id teacher_prefix school_state project_submitted_datetime  \
55966  p205746           Ms.           UT        2016-08-25 01:29:34
55531  p042481          Mrs.           AL        2017-01-25 21:29:09
```

```
26116  p093362            Mr.           CA        2017-01-10 13:35:13
46770  p026359           Mrs.           OK        2016-09-14 16:01:00
80292  p134644           Mrs.           LA        2016-12-11 15:40:18

      project_grade_category          project_subject_categories  \
55966            Grades_9_12                  History & Civics
55531             Grades_3_5                 Literacy & Language
26116             Grades_3_5  Math & Science, Literacy & Language
46770             Grades_3_5     Applied Learning, Special Needs
80292          Grades_PreK_2  Literacy & Language, Math & Science

              project_subject_subcategories  \
55966    History & Geography, Social Sciences
55531         Literacy, Literature & Writing
26116  Applied Sciences, Literature & Writing
46770       Early Development, Special Needs
80292                  Literacy, Mathematics

                            project_title  \
55966        hold tight using ipad case grip
55531     learning with the click of a mouse
26116                   technology title 1
46770  speech language special needs children
80292                         techi firsties

                            project_resource_summary  \
55966  my students need ipad case drop new ipad not b...
55531  my students need benefit using usb wired mice ...
26116  my students need chromebooks research create c...
46770  my students need ipad air 2 access speech lang...
80292  my students need 10 samsung galaxy tablets 10 ...

       teacher_number_of_previously_posted_projects  project_is_approved  \
55966                                             2                   1
55531                                             0                   1
26116                                             2                   1
46770                                             0                   1
80292                                             0                   1

                                                   essay   price  quantity
55966  our high school school residing salt lake city...  102.31         2
55531  my school encompasses k 4th grade students we ...    6.29        26
26116  our school one oldest district many students c...  175.42         5
46770  i serve students special needs the area live h...  551.23         2
80292  we young energetic group first graders come di...  650.92        30
```



**********5 cluster**********

	id	teacher_prefix	school_state	project_submitted_datetime
69262	p013133	Mrs.	NY	2016-10-15 16:53:46
59836	p056013	Mrs.	TX	2017-01-03 22:43:46
10776	p062274	Ms.	KY	2017-03-28 18:22:57
10240	p008999	Mr.	MA	2016-08-14 00:00:23
274	p006577	Ms.	IL	2016-07-25 17:14:02

	project_grade_category	project_subject_categories
69262	Grades_3_5	Special Needs
59836	Grades_PreK_2	Applied Learning, Literacy & Language
10776	Grades_PreK_2	Literacy & Language, Math & Science
10240	Grades_3_5	Health & Sports
274	Grades_9_12	Math & Science

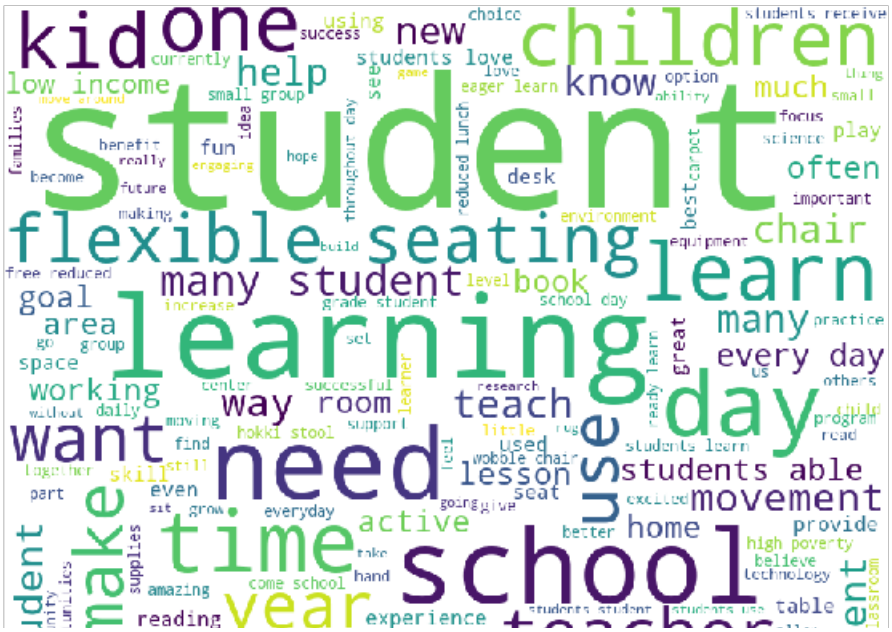
	project_subject_subcategories
69262	Special Needs
59836	Early Development, Literacy
10776	Literacy, Mathematics
10240	Gym & Fitness, Health & Wellness
274	Environmental Science, Health & Life Science

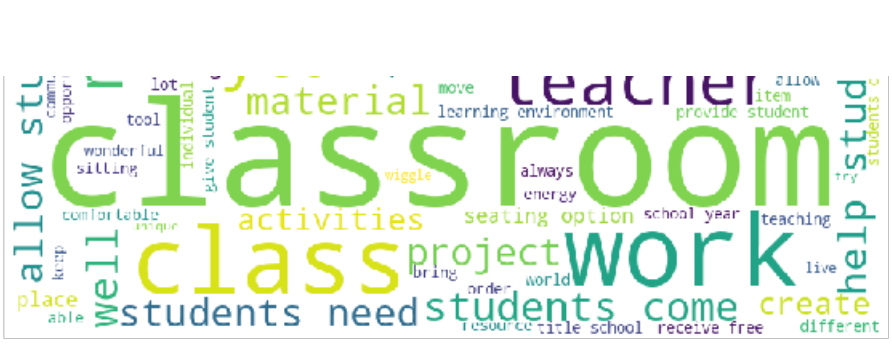
	project_title
69262	wobbly learning
59836	wiggle wiggle wiggle
10776	wobbling working
10240	warming up with fitness gaming
274	classroom pets fish tadpoles turtles chameleon...

	project_resource_summary
69262	my students need five wobble chair they energy...
59836	my students need hokki stools help get wiggles...
10776	my students need 5 wobble stools replace yoga ...
10240	my students need playstation 4 playstation cam...
274	my students need goldfish tropical fish tadpol...

	teacher_number_of_previously_posted_projects	project_is_approved
69262	20	0
59836	18	1
10776	5	1
10240	74	1
274	49	1

	essay	price	quantity
69262	my students need six learniture active learnin...	209.85	5
59836	i blessed teach fun loving group five year old...	392.03	9
10776	our classroom helps get wiggles i work this co...	50.12	5
10240	what learning today that something i hear stud...	672.03	13
274	my students freshmen co taught biology classro...	425.16	49





**********6 cluster**********

	id	teacher_prefix	school_state	project_submitted_datetime	\
13881	p204466	Ms.	LA	2016-12-06 13:27:05	
24699	p117679	Ms.	NY	2017-03-18 22:59:26	
104255	p179350	Mr.	IN	2016-08-25 21:26:08	
98081	p169306	Ms.	CO	2017-01-08 12:25:35	
47132	p142159	Mrs.	MS	2016-05-25 14:35:40	

	project_grade_category	project_subject_categories	\
13881	Grades_3_5	Literacy & Language, Music & The Arts	
24699	Grades_3_5	Applied Learning, Music & The Arts	
104255	Grades_PreK_2	Music & The Arts	
98081	Grades_9_12	Music & The Arts	
47132	Grades_PreK_2	Music & The Arts	

	project_subject_subcategories	\
13881	Literature & Writing, Music	
24699	Character Education, Music	
104255	Music	
98081	Music, Performing Arts	
47132	Music, Performing Arts	

	project_title	\
13881	as mama cass sang make your own kind music	
24699	come shake body baby conga	
104255	sensory instruments just make sense	
98081	help us stand out rehearsal	
47132	meaningful music	

	project_resource_summary	\
13881	my students need apple macbook pro laptop help...	
24699	my students need congas complete latin rhythm ...	
104255	my students need sensory instruments explore n...	
98081	my students need way move stands classrooms no...	
47132	my students need musical cd like casper baby p...	

	teacher_number_of_previously_posted_projects	project_is_approved	\
13881	8	1	
24699	2	1	
104255	1	0	
98081	0	1	
47132	1	1	

	essay	price	quantity
13881	our pre k 5th grade school one elementary scho...	1499.99	1
24699	vivir mi vida la la la we growing modern ba...	269.99	1
104255	my school kindergarten sixth grade environment...	179.63	15
98081	our school growing program three orchestras co...	305.86	1
47132	my students love explore world eyes wide open ...	108.46	8

**********7 cluster**********

	id	teacher_prefix	school_state	project_submitted_datetime
48084	p011894	Ms.	WI	2016-07-18 14:43:03
59452	p226149	Mrs.	ME	2017-03-07 13:01:09
93349	p236406	Ms.	ME	2016-09-19 22:01:59
26826	p185758	Ms.	KY	2016-11-14 19:57:32
84107	p214781	Mrs.	SC	2016-09-19 20:26:27

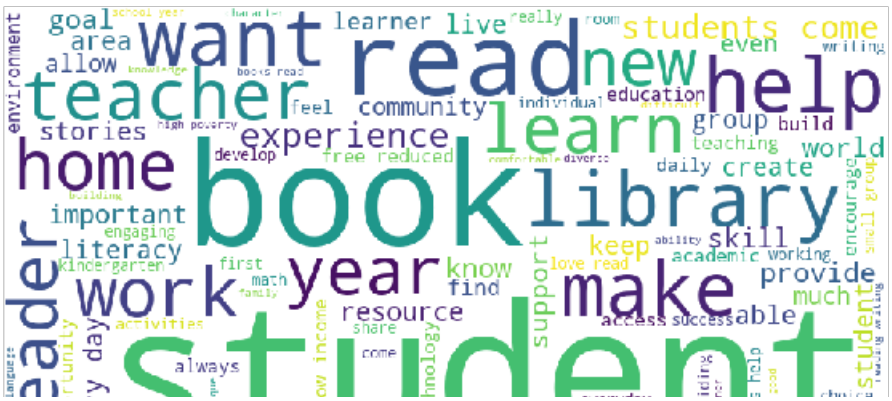
	project_grade_category	project_subject_categories
48084	Grades_3_5	Literacy & Language
59452	Grades_PreK_2	Literacy & Language
93349	Grades_6_8	Literacy & Language
26826	Grades_3_5	Literacy & Language
84107	Grades_PreK_2	Literacy & Language

	project_subject_subcategories	project_title
48084	Literacy	third grade book lovers
59452	Literacy	spruce seuss collection
93349	Literacy, Literature & Writing	read your heart out with us
26826	Literacy	student choice means student growth
84107	Literacy	the just right library

	project_resource_summary
48084	my students need variety texts grow independen...
59452	my students need dr seuss books add learning l...
93349	my students need hands books authors hopefully...
26826	my students need storage individualized differ...
84107	my students need leveled books support literac...

	teacher_number_of_previously_posted_projects	project_is_approved
48084	0	1
59452	21	1
93349	1	1
26826	1	0
84107	2	1

	essay	price	quantity
48084	i entering classroom wonderful unique third gr...	455.14	9
59452	my students come various backgrounds we teach ...	192.94	28
93349	do love read heart we in february school host ...	20.83	21
26826	the students fifth grade classroom come divers...	64.48	4
84107	in kindergarten students come many different a...	461.43	3



5. Applying DBSCAN

In [13]:

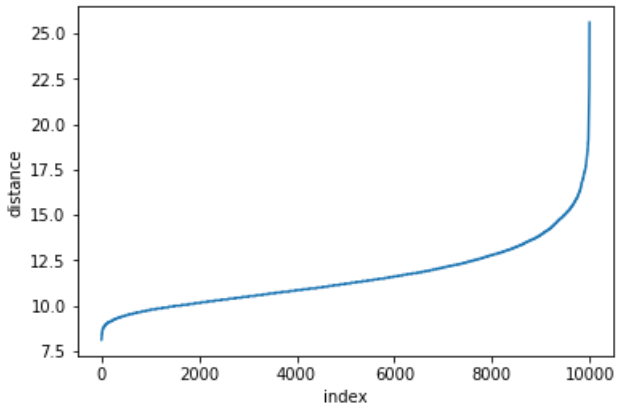
```
from sklearn.neighbors import NearestNeighbors
nbrs=NearestNeighbors(100)
nbrs.fit(x_train_new)
distances, indices = nbrs.kneighbors(x_train_new)
```

In [14]:

```
plt.plot(sorted(distances[:,-1]))
plt.xlabel('index')
plt.ylabel('distance')
```

Out[14]:

```
Text(0, 0.5, 'distance')
```



In [16]:

```
#optimal value of epsilon=15
def infoClusters(num_clusters,predicted):
    for i in range(num_clusters):
        print(("*"*20)+str(i)+" cluster"+("*"*20))
        temp=(predicted==i)
        print(data[temp][:5])
        wordcloud(temp)



dbscan = DBSCAN(min_samples=100,eps=15)
predicted = dbscan.fit_predict(x_train_new.toarray())
num_clusters=np.max(predicted)+1
print(("*"*20)+" DBSCAN Clustering"+("*"*20))
infoClusters(num_clusters,predicted)
```

```
******************** DBSCAN Clustering********************
*********************0 cluster*********************
             id teacher_prefix school_state project_submitted_datetime  \
```

	id	teacher_prefix	school_state	project_submitted_datetime	\
105447	p118457	Teacher	NC	2016-07-05 16:12:10	
69262	p013133	Mrs.	NY	2016-10-15 16:53:46	
55966	p205746	Ms.	UT	2016-08-25 01:29:34	
1860	p115325	Mrs.	OR	2016-10-31 23:30:00	
35390	p158955	Ms.	CA	2016-07-22 16:42:42	

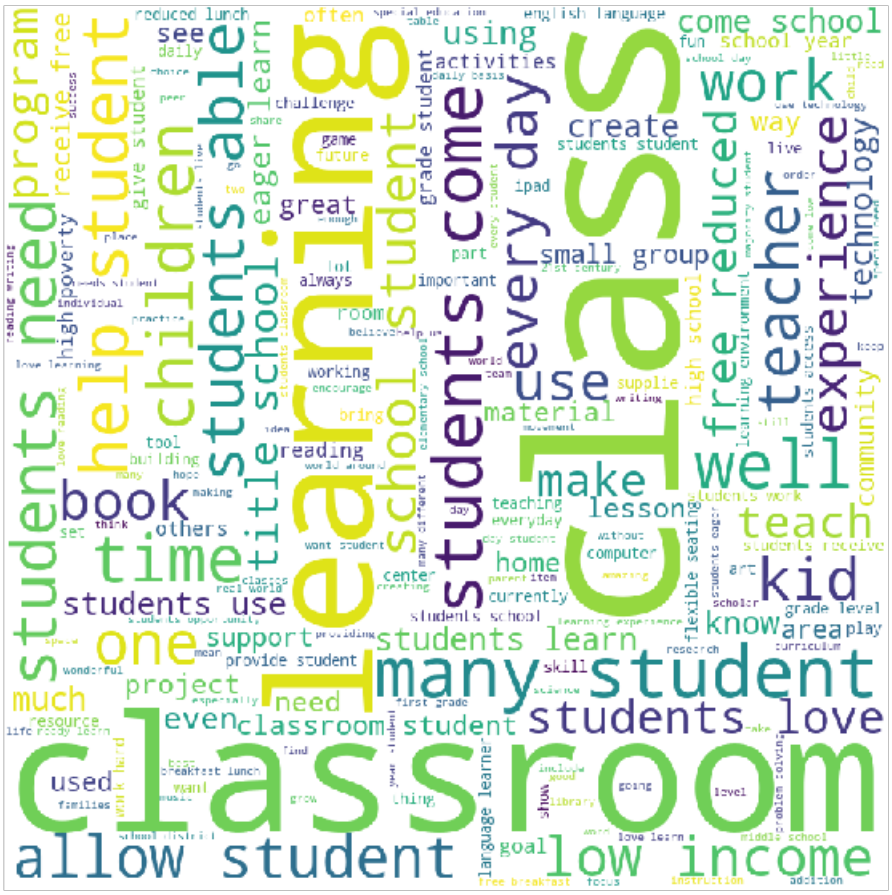
	project_grade_category	project_subject_categories	\
105447	Grades_9_12	Math & Science, Special Needs	
69262	Grades_3_5	Special Needs	
55966	Grades_9_12	History & Civics	
1860	Grades_3_5	Music & The Arts	
35390	Grades_6_8	Math & Science	

	project_subject_subcategories	project_title	\
105447	Environmental Science, Special Needs	rain barrels school garden	
69262	Special Needs	wobbly learning	
55966	History & Geography, Social Sciences	hold tight using ipad case grip	
1860	Visual Arts	clean surfaces watercolor art	
35390	Environmental Science	click engage	

	project_resource_summary	\
105447	my students need rain barrels collect water ga...	
69262	my students need five wobble chair they energy...	
55966	my students need ipad case drop new ipad not b...	
1860	my students need watercolor paper art cleaning...	
35390	my students need 34 interactive clickers make ...	

	teacher_number_of_previously_posted_projects	project_is_approved	\
105447	0	1	
69262	20	0	
55966	2	1	
1860	1	1	
35390	0	1	

	essay	price	quantity
105447	environmental science students no idea food co...	93.69	4
69262	my students need six learniture active learnin...	209.85	5
55966	our high school school residing salt lake city...	102.31	2
1860	my students love take pride rural community th...	50.07	8
35390	my middle school located underserved community...	1541.14	3



6. Conclusions

K Means:

- K means is doing good job in seperating data points based on characterstics.And we can observe project categories are grouped as clusters.

Agglomerative:

- Agglomerative clustering with number of clusetrs as 2 is not able to cluster the data effectively.But Where as with number of cluster as 4 it was doing decent job.

DBSCAN:

- Unfortunately DBSCAN algorithm is not able to good job as it gave everthing as one cluster.one possible reason could be dataset is highly overlapped in 5k dimension space.