

Analisis de Encuestas Version Publica

July 16, 2017

1 ENCUESTA DE CURSADA

1.1 Primer Cuatrimestre 2017

1.1.1 Departamento de Computación FIUBA

```
In [30]: import numpy as np
import pandas as pd
# plots
import matplotlib.pyplot as plt
import seaborn as sns
%matplotlib inline
import sys
reload(sys)
sys.setdefaultencoding('utf8')
```

```
In [8]: comp = pd.read_csv('encuesta2.csv')
comp.sample(1)
```

```
Out[8]:
```

	Timestamp	Curso	\
318	2017/07/03 1:45:01 PM GMT-3	Base de Datos - Beiro	
	Opinión General Sobre el Curso ¿Aprobó la Cursada?		\
318	Muy Bueno	No	
	¿Cómo te Resultaron los Temas de la Materia?		\
318	Interesantes		
	¿Los Temas de la Materia Están Actualizados?		\
318	Sí		
	Nivel de las Clases Teóricas del Curso		\
318	Excelentes		
	Nivel de las Clases Prácticas del Curso	Dificultad del Curso	\
318	Muy Buenas	Muy Dificil	
	Dificultad del TP Comentarios Sobre el Curso		
318	Dificil	-	

Total de Encuestas Procesadas

```
In [9]: len(comp)
```

```
Out[9]: 528
```

```
In [10]: # This not used :)
```

```
filter = (comp["Comentarios Sobre el Curso"].notnull())
```

```
opinions = comp.loc[filter,['Curso','Comentarios Sobre el Curso']].sort_va
```

```
In [11]: map_sino = {'Sí':5,'No':1}
```

```
map_opinion = {'Excelente':5,'Muy Bueno':4, 'Bueno':3,'Regular':2,'Malo':1}
```

```
map_temas = {'Interesantes':4,'Muy Interesantes':5, 'Poco Interesantes':2,
```

```
map_clases_teoricas = {'Excelentes':5,'Muy Buenas':4, 'Buenas':3,'Regulare
```

```
map_dificultad_tp = {'Muy dificil':5,'Dificil':4,'Normal':3,'Fácil':2,'Muy
```

```
map_dificultad = {'Muy Dificil':5,'Dificil':4,'Normal':3,'Facil':2,'Muy Fa
```

```
map_clases_practicas = {'Excelentes':5,'Muy Buenas':4, 'Buenas':3,'Regular
```

```
comp.replace({"¿Aprobó la Cursada?": map_sino},inplace=True)
```

```
comp.replace({"¿Los Temas de la Materia Están Actualizados?": map_sino},in
```

```
comp.replace({"Opinión General Sobre el Curso": map_opinion},inplace=True)
```

```
comp.replace({"Nivel de las Clases Teóricas del Curso": map_clases_teorica
```

```
comp.replace({"Nivel de las Clases Prácticas del Curso": map_clases_practi
```

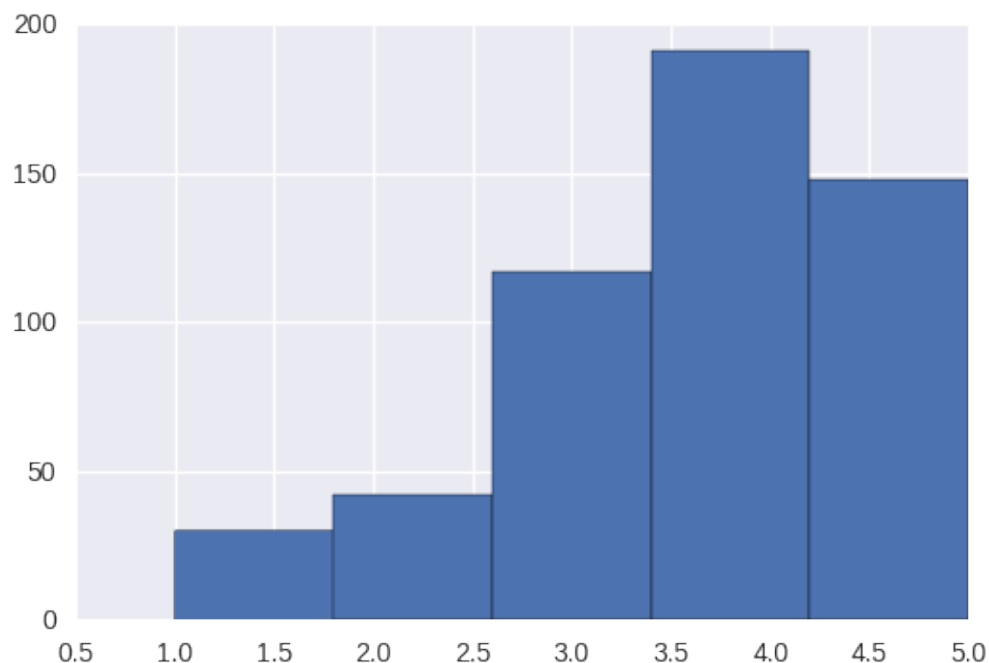
```
comp.replace({"Dificultad del Curso": map_dificultad},inplace=True)
```

```
comp.replace({"Dificultad del TP": map_dificultad_tp},inplace=True)
```

```
comp.replace({"¿Cómo te Resultaron los Temas de la Materia?": map_temas},i
```

1.2 Histograma de calificaciones para la opinión general sobre el curso

```
In [12]: comp['Opinión General Sobre el Curso'].hist(bins=5);
```

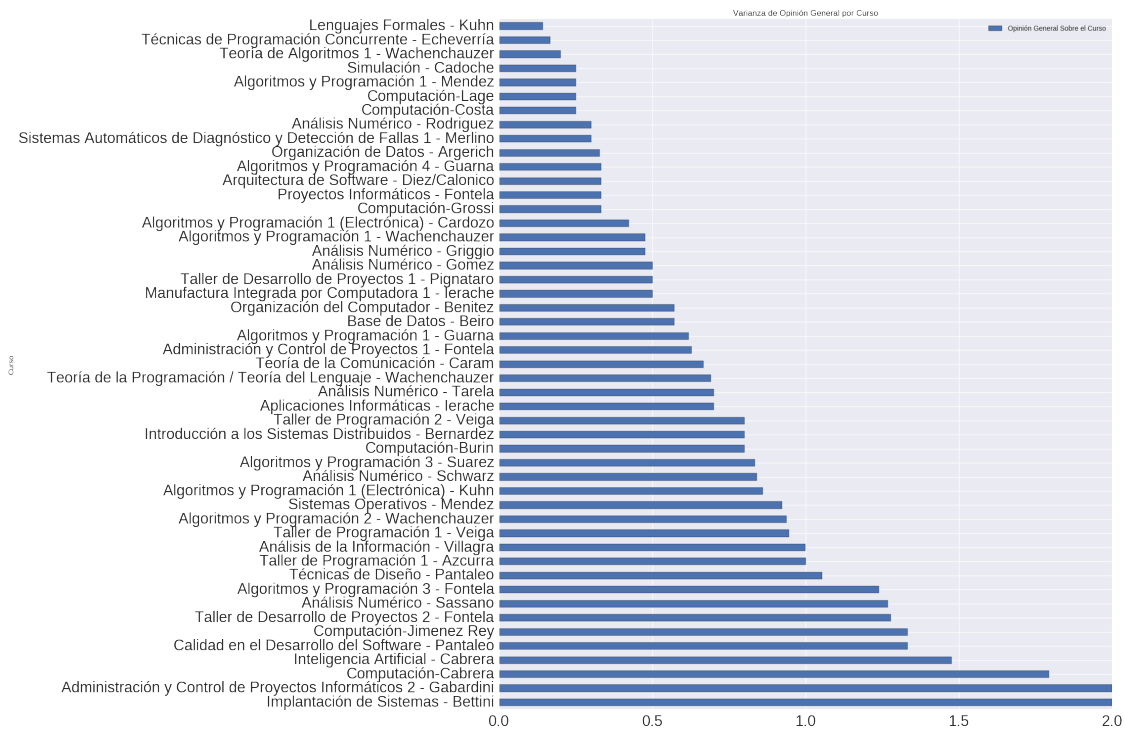


```
In [13]: grouped = comp.groupby('Curso')
num_votes = grouped.agg('size')
g_mean = grouped.mean()
g_mean['votes']=num_votes
```

1.3 Varianza de la Opinión por Curso

Este gráfico muestra la variabilidad de la pregunta “Opinión General Sobre el Curso”, a menor valor mas consistentes los resultados de la encuesta. Los valores mayores muestran disparidad de resultados en cuanto a la opinión general sobre el curso.

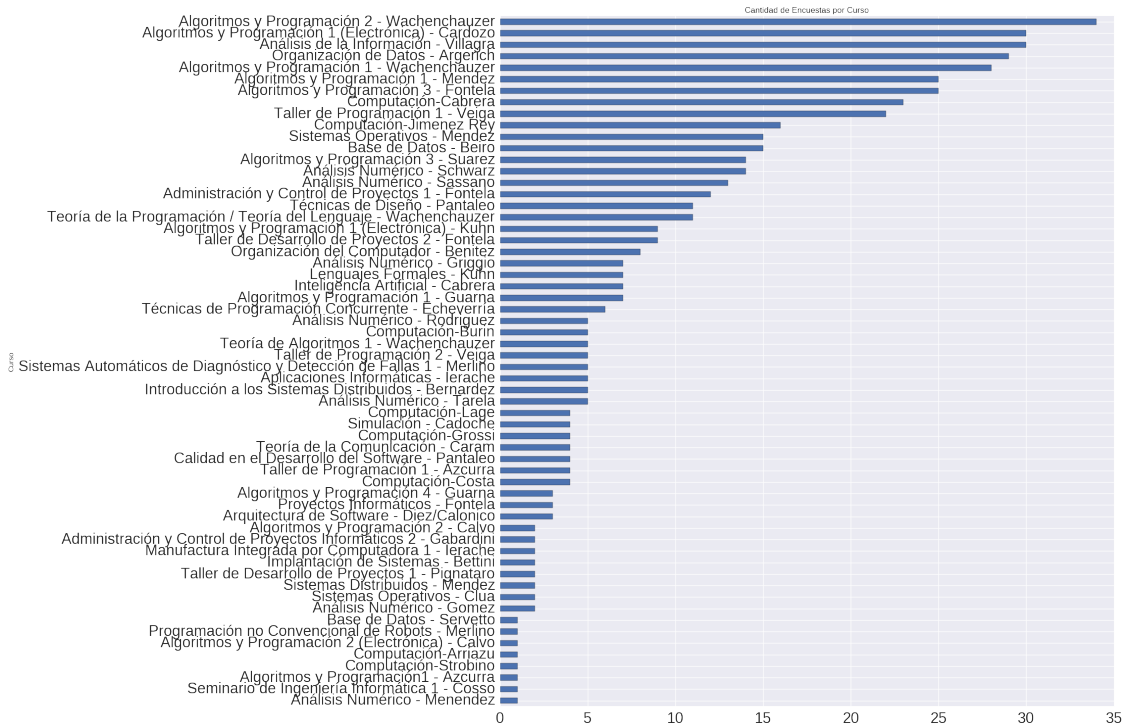
```
In [14]: g_var = grouped.var()
gop = g_var.loc[g_var['Opinión General Sobre el Curso']>0,['Opinión General Sobre el Curso']
sort_values(by='Opinión General Sobre el Curso',ascending=False).\
plot(kind='barh',figsize=(16,18),title='Varianza de Opinión General por Curso')
```



1.4 Cantidad de Encuestados por Curso

Cantidad de encuestas completadas por cada Curso.

```
In [15]: g_mean['votes'].sort_values(ascending=True).plot(kind='barh',figsize=(16,18))
```

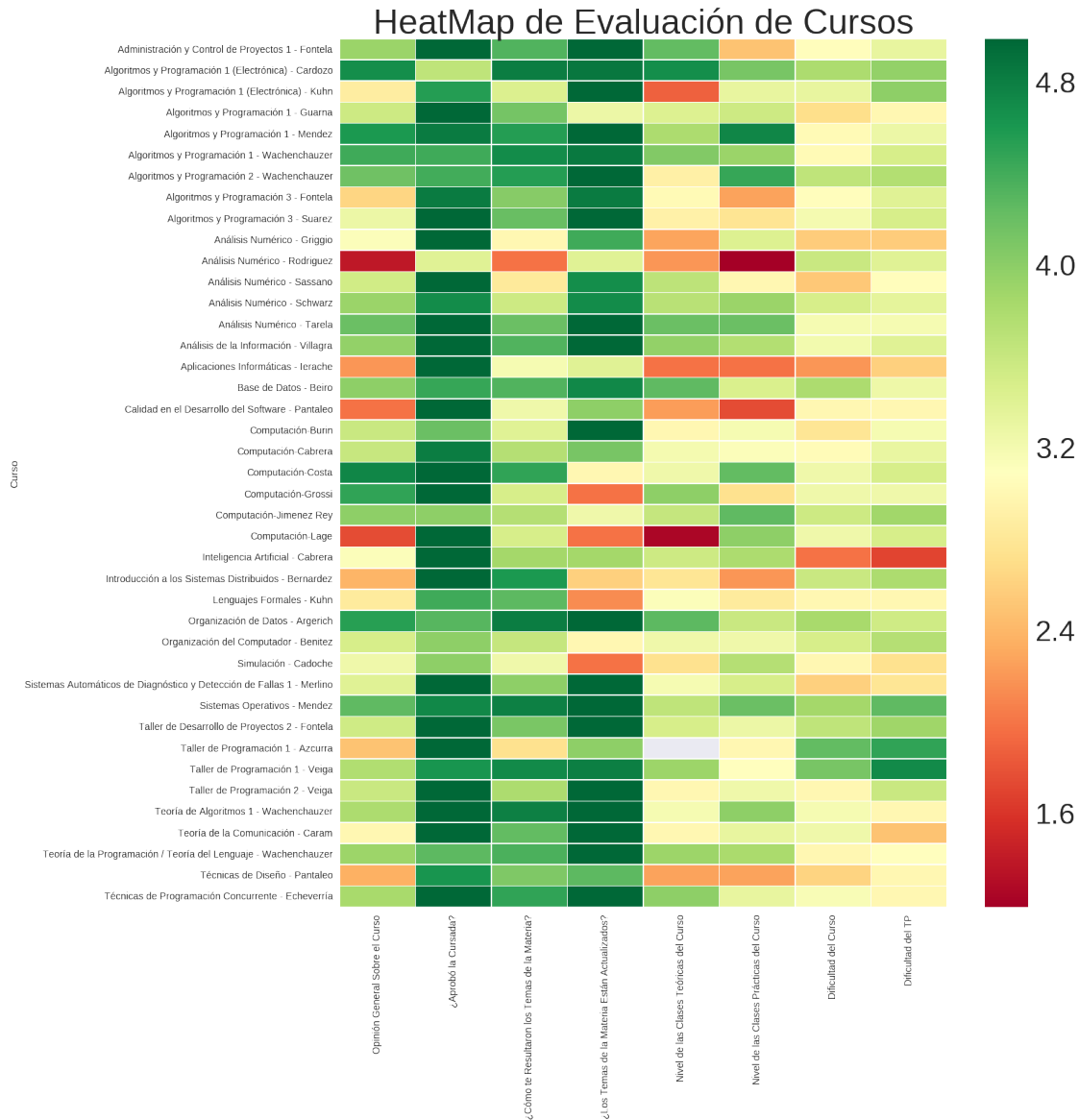


```
In [16]: # Quedarnos unicamente con los cursos que tienen 4 o mas encuestas
g_mean = g_mean.loc[g_mean['votes']>3,:]
lista_cursos = g_mean.index.tolist()
```

1.5 Heatmap de Evaluación de Cursos (5 o mas respuestas)

Esta matriz muestra el promedio de calificaciones para cada pregunta de la encuesta para los cursos que tienen 5 o mas encuestas. Orden alfabético por curso.

```
In [17]: g_mean2 = g_mean.drop('votes',1)
fig, ax = plt.subplots(figsize=(14,16))
sns.set(font_scale=3.0)
sns.heatmap(g_mean2, annot=False, linewidths=.5, cmap='RdYlGn', ax=ax);
plt.title('HeatMap de Evaluación de Cursos');
```

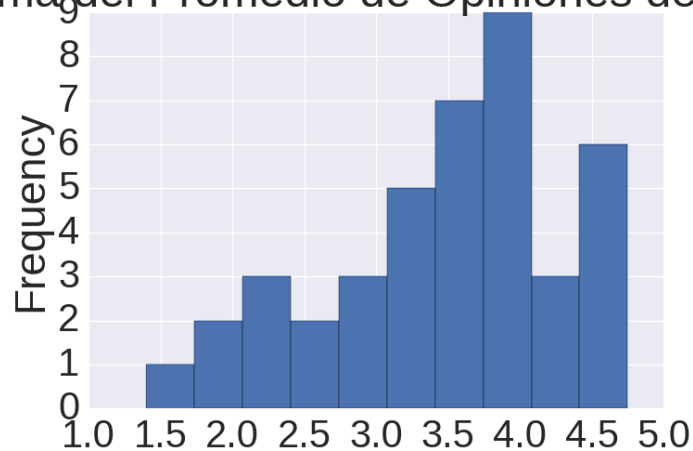


1.6 Histograma del Promedio de Opiniones de los Cursos (un promedio x curso)

Por cada curso se calcula el promedio de la pregunta “Opinión general sobre el curso” y luego se hace un histograma de los resultados. El gráfico solo incluye aquellos cursos con 5 o mas encuestas.

```
In [18]: g_mean['Opinión General Sobre el Curso'].plot(kind='hist',bins=10,title='P
```

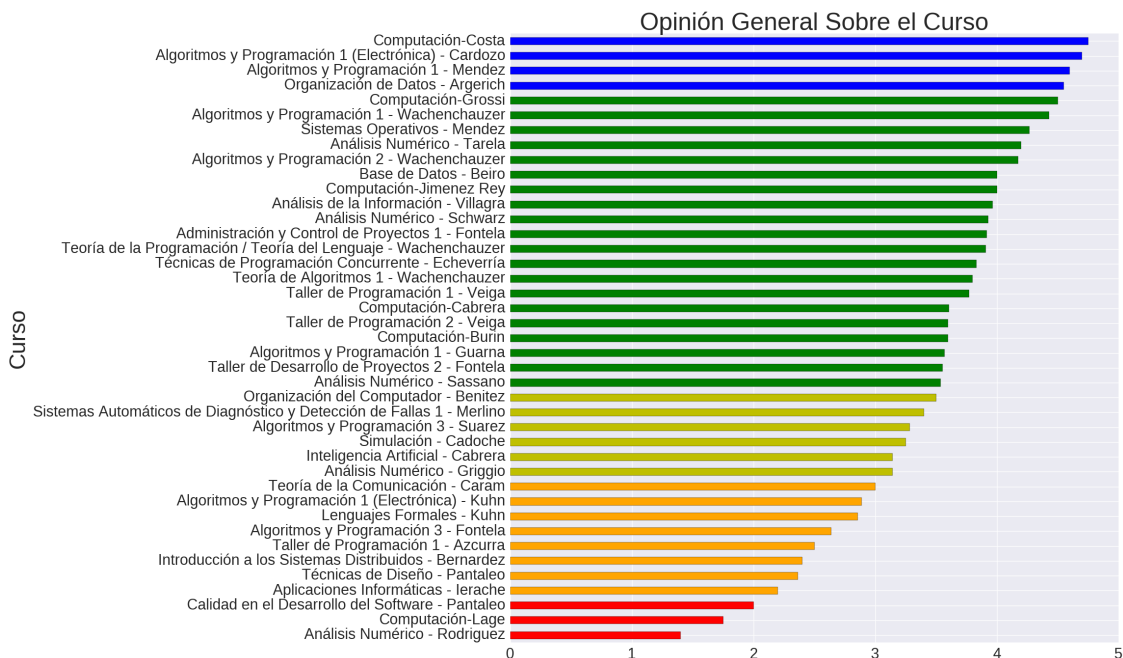
Histograma del Promedio de Opiniones de los Cursos



1.7 Opinión General Sobre el Curso

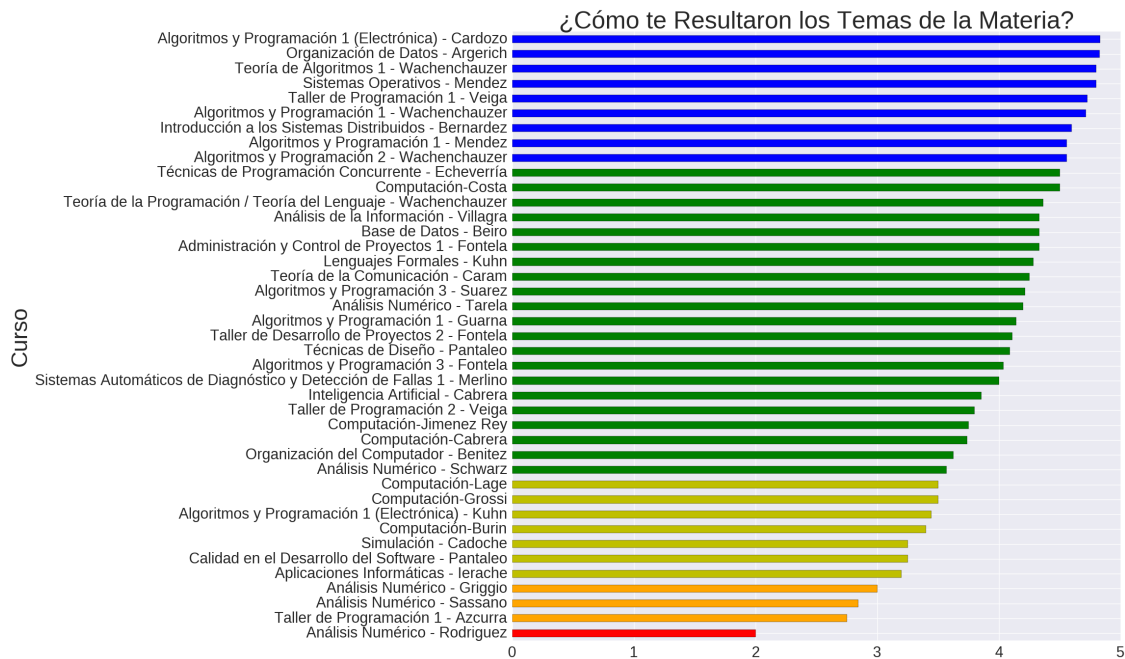
```
In [19]: CUTLIST = [0, 2, 3, 3.5, 4.5, 5]
        LABELSLIST = ['r', 'orange', 'y', 'g', 'b']
```

```
In [20]: VARIABLE = 'Opinión General Sobre el Curso'
        g_mean = g_mean.sort_values(by=VARIABLE, ascending=True)
        colors = pd.cut(g_mean[VARIABLE], CUTLIST, labels=LABELSLIST)
        g_mean[VARIABLE].plot(kind='barh', color= colors, figsize=(16,16), fontsize=20)
```



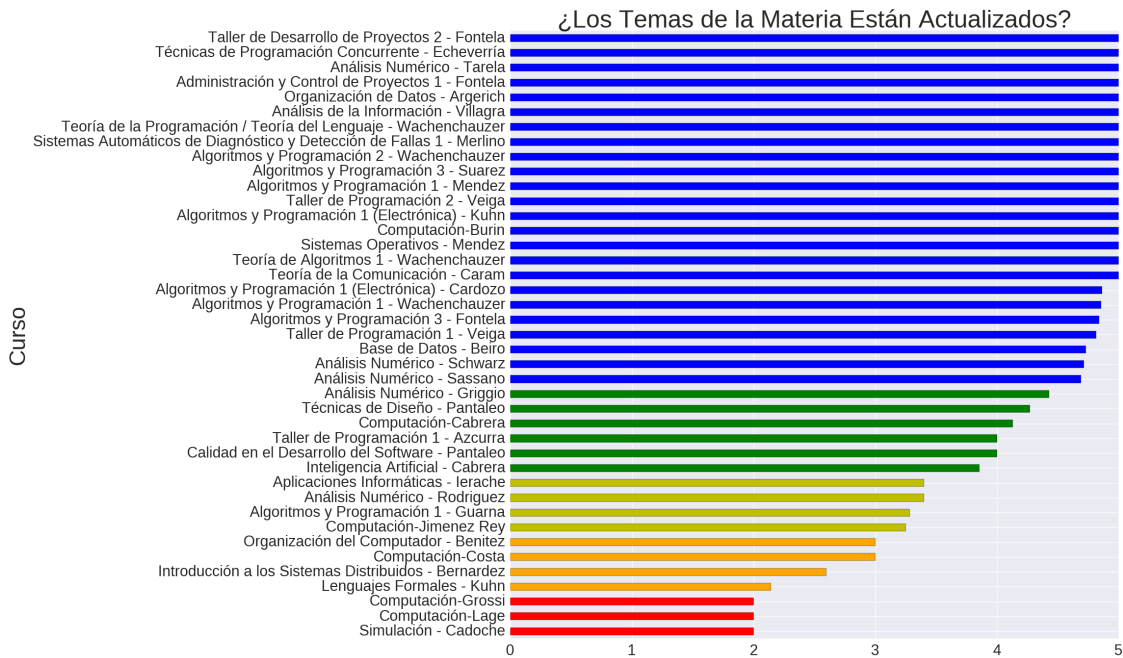
1.8 ¿Los Temas de la Materia son Interesantes?

```
In [21]: VARIABLE = '¿Cómo te Resultaron los Temas de la Materia?'
g_mean = g_mean.sort_values(by=VARIABLE, ascending=True)
colors= pd.cut(g_mean[VARIABLE], CUTLIST, labels=LABELSLIST)
g_mean[VARIABLE].plot(kind='barh', color= colors, figsize=(16,16), fontsize=20)
```



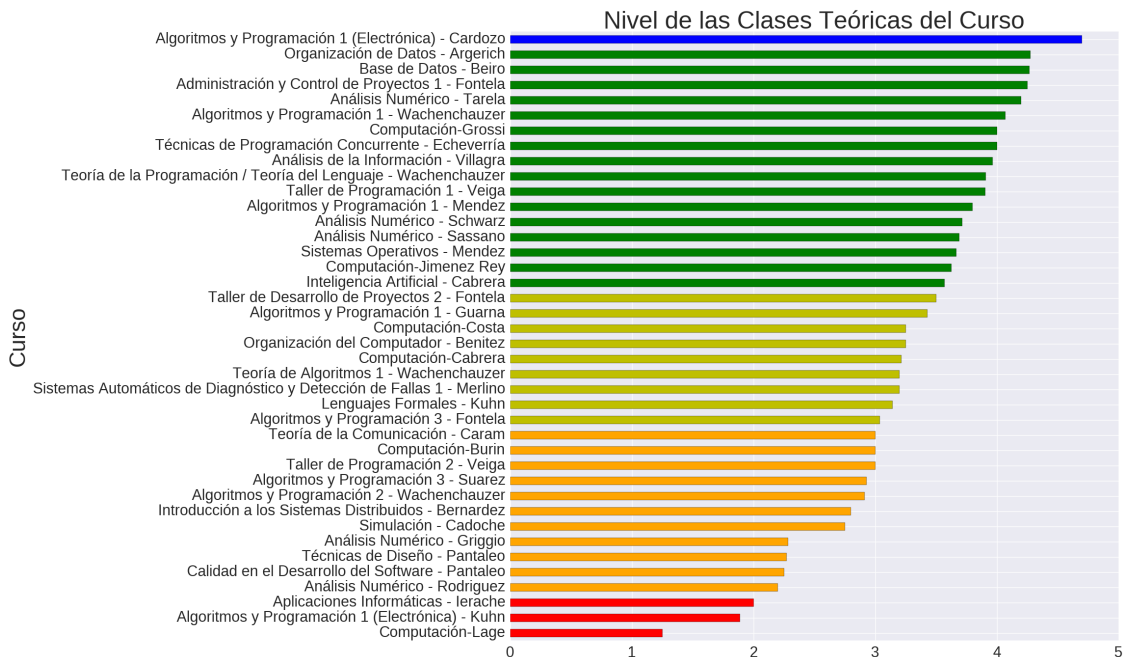
1.9 ¿Los Temas Están Actualizados?

```
In [22]: VARIABLE = '¿Los Temas de la Materia Están Actualizados?'
g_mean = g_mean.sort_values(by=VARIABLE, ascending=True)
colors = pd.cut(g_mean[VARIABLE], CUTLIST, labels=LABELSLIST)
g_mean[VARIABLE].plot(kind='barh', color= colors, figsize=(16,16), fontsize=20)
```



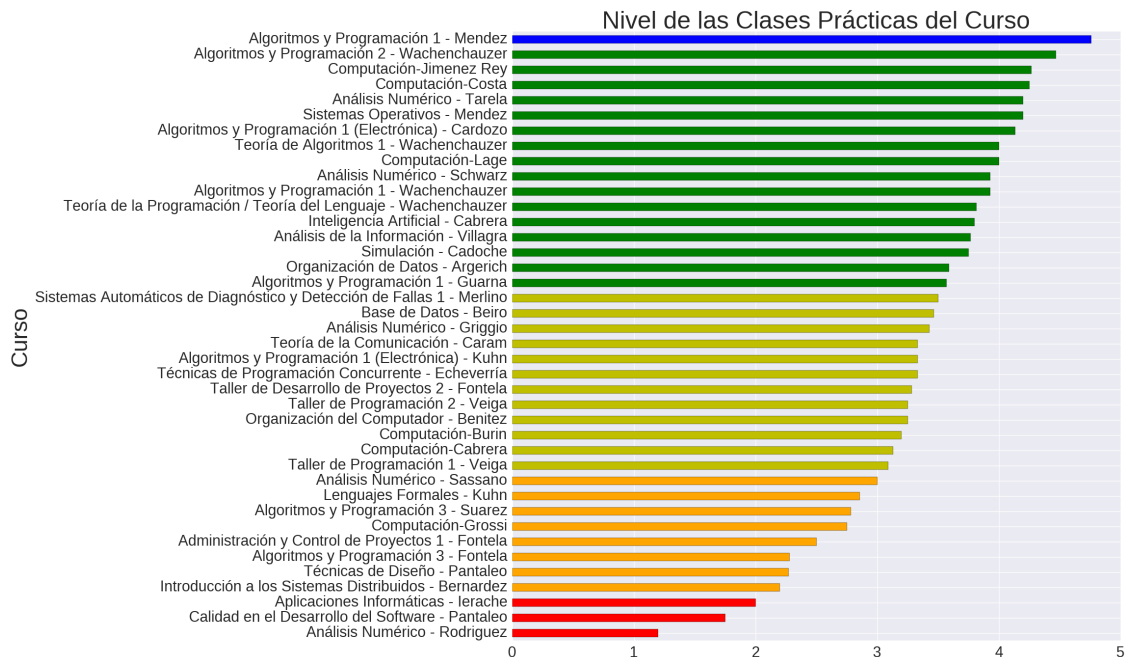
1.10 Nivel de las Clases Teóricas

```
In [23]: VARIABLE = 'Nivel de las Clases Teóricas del Curso'
g_mean = g_mean.sort_values(by=VARIABLE,ascending=True).dropna()
colors = pd.cut(g_mean[VARIABLE], CUTLIST, labels=LABELSLIST)
g_mean[VARIABLE].plot(kind='barh',color= colors,figsize=(16,16),fontsize=20)
```



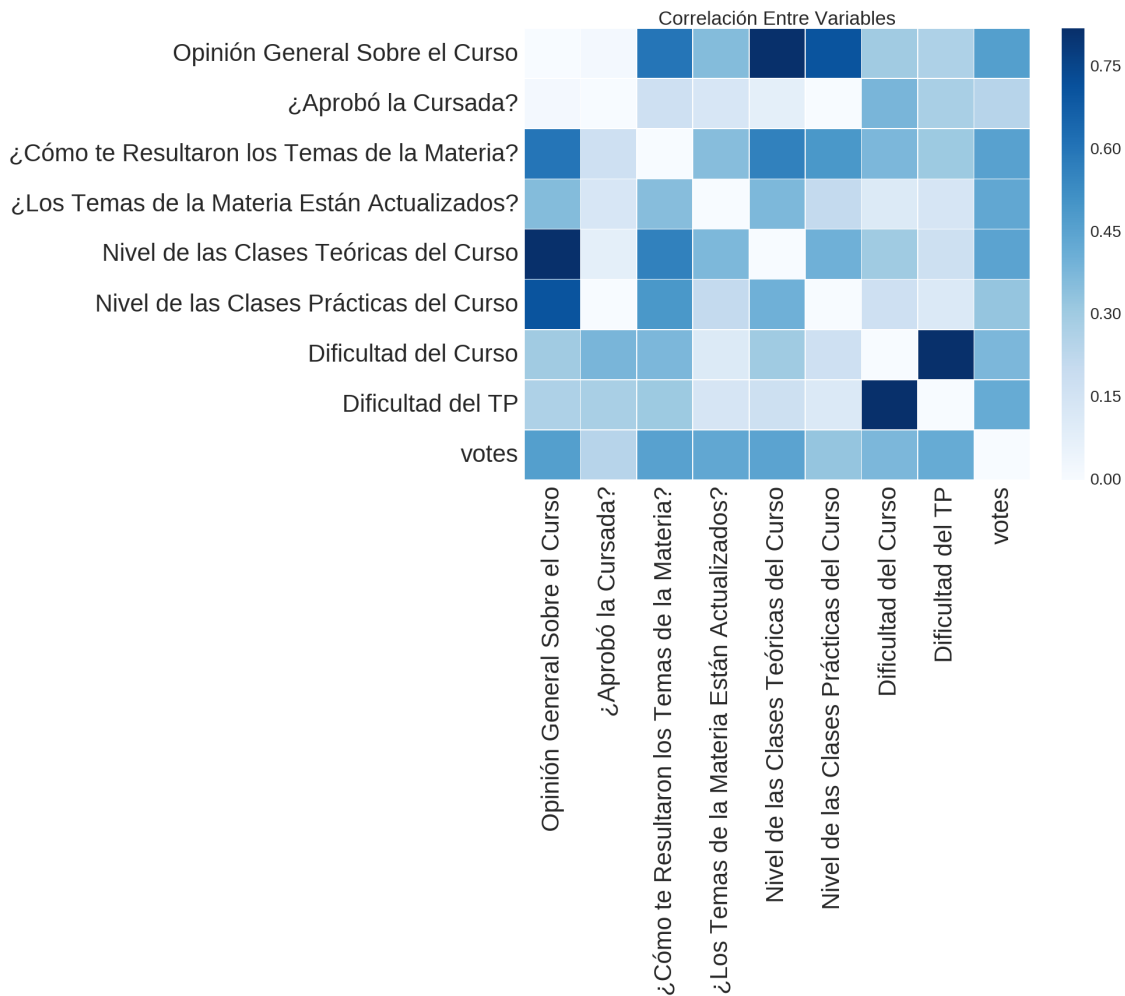
1.11 Nivel de las Clases Prácticas del Curso

```
In [24]: VARIABLE = 'Nivel de las Clases Prácticas del Curso'
g_mean = g_mean.sort_values(by=VARIABLE, ascending=True).dropna()
colors = pd.cut(g_mean[VARIABLE], CUTLIST, labels=LABELSLIST)
g_mean[VARIABLE].plot(kind='barh', color= colors, figsize=(16,16), fontsize=20)
```



1.12 Correlacion Entre Variables

```
In [25]: cor = g_mean.corr()
cor.values[np.arange(9)*2] = 0
cor.drop('votes', axis=1)
fig, ax = plt.subplots(figsize=(14,10)) # Sample figsize in inches
sns.set(font_scale=2)
sns.heatmap(cor.applymap(abs), annot=False, linewidths=.5, cmap='Blues', ax=ax)
plt.title('Correlación Entre Variables');
```

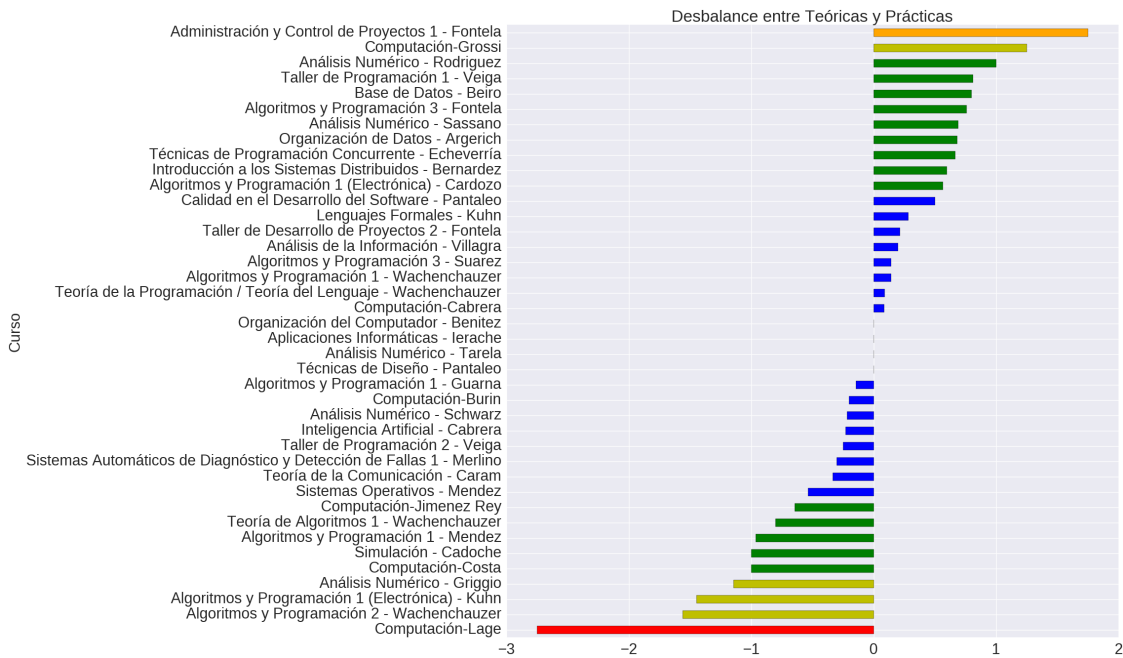


1.13 Diferencia Entre Nivel de Clases Teóricas y Prácticas

Hacia la derecha: Las clases teóricas son mejores que las prácticas. Hacia la izquierda: Las clases prácticas son mejores que las teóricas. En el centro: Poca o Ninguna diferencia.

```
In [26]: g_mean['dif_teo_prac'] = ((g_mean['Nivel de las Clases Teóricas del Curso'] -
g_mean['Nivel de las Clases Prácticas del Curso']) / 2)

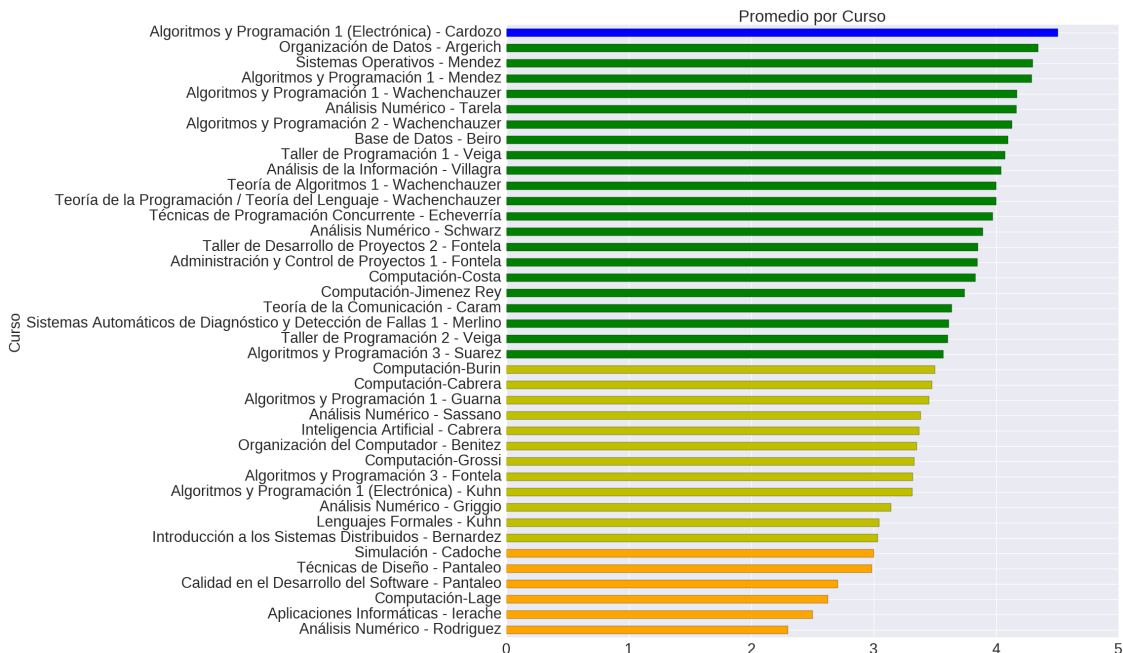
In [27]: VARIABLE = 'dif_teo_prac'
g_mean = g_mean.sort_values(by=VARIABLE, ascending=True).dropna()
colors = pd.cut(-abs(g_mean[VARIABLE]), 5, labels=LABELSLIST)
g_mean[VARIABLE].plot(kind='barh', color= colors, figsize=(16,16), fontsize=20)
```



1.14 Promedio General por Curso

```
In [28]: g_mean['general_ave'] = (g_mean['Opinión General Sobre el Curso']+g_mean['
```

```
In [29]: VARIABLE = 'general_ave'
g_mean = g_mean.sort_values(by=VARIABLE,ascending=True).dropna()
colors = pd.cut(g_mean[VARIABLE], CUTLIST, labels=LABELSLIST)
g_mean[VARIABLE].plot(kind='barh',color= colors,figsize=(16,16),fontsize=2
```



In []: