	База данных для функционирования института Описание бд
	Данная предметная область представляет собой институт, например МГТУ им. Баумана, в котором работают и обучаются множество людей. У каждого сотрудника или студента есть своя информация, необходимая для обучения или работы в данном учреждении. А также у каждого человека в данной организации существует информация, которая формируется с течением времени. Например, стипендия, курс, семестр студента. Соответственно, база данных реализованная для этой области поможет структурировать информацию данной организации и упростит работу и отслеживания изменений. Таблицы бд
	1. Departament - Таблица, содержащая институты 2. Student_Counceling - Таблица, содержащая общую информацию студентов обучающихся в институтах 3. Student_Perfomance - Таблица, содержащая информацию успеваемости студентов
	Student_Counceling I Student_D DOA DOB Department_Choices Department_Admission Department_Name DOE Employee Employee Employee DOB
In [62]:	4. Employee - Таблица, содержащая общую информацию преподователей преподающих в институтах ### Имфологическая модель import numpy as np import pandas as pd
	<pre>import pandas as pd import seaborn as sns import matplotlib.pyplot as plt %matplotlib inline sns.set(style="ticks")</pre>
In [64]:	<pre>Depart = pd.read_csv('Departament.csv', sep=";") StCoun = pd.read_csv('Student_Counceling.csv', sep=";") StPerf = pd.read_csv('Student_Performance.csv', sep=";") Employee = pd.read_csv('Employee.csv', sep=";")</pre>
In [65]:	print("Размер таблицы Departament: ", Depart.shape) print("Размер таблицы Student_Counceling: ", StCoun.shape) print("Размер таблицы Student_Performance: ", StPerf.shape) print("Размер таблицы Employee: ", Employee.shape) Размер таблицы Departament: (40, 3)
In [66]:	Pазмер таблицы Student_Counceling: (3819, 5) Pазмер таблицы Student_Performance: (209611, 6) Pазмер таблицы Employee: (998, 4) print("Departament:") Depart.head()
Out[66]:	Department ID Department Name Dopartment Name <th< td=""></th<>
In [67]:	Student_Counceling: Student_Counceling:
Out[67]: In [68]:	0 SID20131143 01.07.2013 05.02.1996 IDEPT783 IDEPT783 1 SID20131151 01.07.2013 31.07.1995 IDEPT6347 IDEPT6347 2 SID20131171 01.07.2013 05.09.1995 IDEPT1836 IDEPT1836 3 SID20131176 01.07.2013 12.01.1996 IDEPT8473 IDEPT8473 4 SID20131177 01.07.2013 30.07.1995 IDEPT5528 IDEPT5528
Out[68]:	Student_Performance: ID Student_ID Semster_Name Paper_ID Paper_Name Marks 0 1 SID20131143 Sem_1 SEMI0012995 Paper 1 44
	1 2 SID20131143 Sem_1 SEMI0015183 Paper 2 74 2 3 SID20131143 Sem_1 SEMI0018371 Paper 3 80 3 4 SID20131143 Sem_1 SEMI0015910 Paper 4 44 4 5 SID20131143 Sem_1 SEMI0016208 Paper 5 95
<pre>In [69]: Out[69]:</pre>	<pre>print("Employee:") Employee.head() Employee:</pre>
In [70]:	0 IU112687 18.06.1982 03.09.2012 IDEPT5127 1 IU114595 20.08.1986 13.11.2001 IDEPT4055 2 IU114836 24.03.1985 23.09.2007 IDEPT601 3 IU116135 10.01.1993 19.02.2002 IDEPT4055 4 IU116424 05.11.1975 22.04.2005 IDEPT1423
	<pre>print('Количество пустых ячеек в таблице Department:') for col in Depart.columns: temp_null_count = Depart[Depart[col].isnull()].shape[0] print('{} - {}'.format(col, temp_null_count))</pre> Kоличество пустых ячеек в таблице Department: Department_ID - 0 Department_Name - 0 DOE - 0
	<pre>print('Количество пустых ячеек в таблице Student_Counceling:') for col in StCoun.columns: temp_null_count = StCoun[StCoun[col].isnull()].shape[0] print('{} - {}'.format(col, temp_null_count))</pre> Kоличество пустых ячеек в таблице Student_Counceling: Student_ID - 0 DOA - 0 DOB - 0 Department_Choices - 0 Department_Admission - 0
	print('Количество пустых ячеек в таблице Student_Performance:') for col in StPerf.columns: temp_null_count = StPerf[StPerf[col].isnull()].shape[0] print('{} - {}^*.format(col, temp_null_count)) Количество пустых ячеек в таблице Student_Performance: ID - 0 Student_ID - 0 Semster_Name - 0 Paper_ID - 0 Paper_Name - 0 Marks - 0
In [73]:	<pre>print('Количество пустых ячеек в таблице Employee:') for col in Employee.columns: temp_null_count = Employee[Employee[Employee[Employee[Employee]]] print('{} - {}'.format(col, temp_null_count))</pre> Kоличество пустых ячеек в таблице Employee: Employee ID - 0 DOB - 0 DOB - 0 Department_ID - 0
<pre>In [74]: Out[74]:</pre>	Department_ID Department_Name DOE
	count4040unique4040topIDEPT1388Centre for Distance Engineering Education Prog22.08.1966freq11
<pre>In [75]: Out[75]:</pre>	StCoun.describe() Student_ID DOA DOB Department_Choices Department_Admission count 3819 3819 3819 3819 3819 unique 3819 6 1759 40 40
In [76]:	top SID20182516 01.07.2015 21.05.1997 IDEPT3115 freq 1 692 9 112 121 # Основные статистические характеристки набора данных StPerf.describe()
	count 299611.00000 209611.00000 mean 104806.00000 695899573 std 60509.627975 18.098626 min 1.00000 19.000000 55% 104806.00000 70.000000 75% 157208.50000 85.000000 max 209611.000000 100.000000
In [77]: Out[77]:	Employee.describe()
	fig, ax = plt.subplots(figsize=(10,10)) print('Распределение успеваемости студентов во всех вузах:') sns.distplot(StPerf['Marks'])
Out[78]:	Pacпределение успеваемости студентов во всех вузах: <axessubplot:xlabel='marks', ylabel="Density"> 0.0200 —</axessubplot:xlabel='marks',>
	0.0175 - 0.0125 - 0.0125 - 0.0075 - 0.0050 - 0.0055 -
In [79]:	20 40 60 80 100 Marks
Out[79]:	<pre> <axessubplot:xlabel='doe', ylabel="Department_Name"> School of Management - Centre for Aerospace Systems Design and Engineering (CASDE) - Computer Centre (CC) - Computer Centre (CC) - ** ** ** ** ** ** ** ** **</axessubplot:xlabel='doe',></pre>
	Contribute Cech (CC) Industrial Design Centre Centre for Policy Studies (CPS) Centre for Policy Studies (CPS) Computer Science & Engineering Centre for Technology Alternatives for Rural Areas (CTARA) Humanities & Social Science National Centre for Mathematics (NCM) ABC-EDS Research Academy Mechanical Engineering Mechanical Engineering Mechanical Technology Mechani
	Electrical Engineering - Energy Science and Engineering - Physics - Physics - Centre for Enterpreneurship (DSCE) - Centre for Urban Science and Engineering (CUSE) - Earth Sciences - Climatic Studies - Centre for Urban Science and Engineering (CUSE) - Earth Sciences - Climatic Studies - Centre for Excellence in Technology for Internal Scoulty (NGCE) - Aurospace Engineering (CUSE) - Earth Sciences - Climatic Studies - Centre for Excellence in Technology and Science (CRNIS) - Busiciences and Bioengineering - Systems and Corrol Engineering - Systems and Corrol Engineering - Systems and Corrol Engineering - Centre for Research in Nanotechnology and Science (CRNIS) - Centre for Research in Anotechnology in Engineering - Systems and Corrol Engineering - Systems and Corrol Engineering - Centre for Research in Nanotechnology in Corrol Engineering - Metallurgical Engineering & Metallurgical Engineering (SSEE) - National Centre for Aerospace Innovation and Research (NCAR) - Centre for Ibidance Engineering & Metallurgical Engin
In [80]:	07.03.283007.08564.26002.08593.103995.19411.19569.08582.28389.200029.09707.25655.28383.260079.194077.23495.08705.223405.28583.260079.194077.23495.08582.283896.195400.08866.20706.28588.198667.196612.195074.089921.283896.20004.198411.1988 DOE x = {} for i in Depart['Department ID']:
In [81]:	<pre>for i in Depart['Department_ID']: x.update({Employee.loc[Employee['Department_ID'] == i].count()[0] : i}) x = dict(sorted(x.items(), key=lambda item: item[0])) ox = [] oy = []</pre>
In [82]:	<pre>for i in x: ox.append(i) oy.append(x[i]) plt.plot(ox,oy)</pre>
	ріт.тіте ("данная диаграмма позволяет отследить количество сотрудников в каждом вузе по отдельности") ріт. show() Данная диаграмма позволяет отследить количество сотрудников в каждом вузе по отдельности IDEPT5109 — IDEPT4670 —
	IDEPT3115 -
[91]:	рlt.ple(ох, labels-оу) plt.title('Данная диаграмма позволяет отследить количество сотрудников в каждом вузе по отдельности') рlt.show() Данная диаграмма позволяет отследить количество сотрудников в каждом вузе по отдельности DEPT8598
	IDEPT8313 IDEPT2357 IDEPT3115
	<pre>sns.heatmap(StPerf.corr()) <axessubplot:></axessubplot:></pre>

ı Marks

In []: