

# Отчет по Домашнему заданию №01 по предмету Проектный семинар “Python в науке о данных”

Тема работы: “Исследование зависимости роста человека по странам”

Анализ подготовили: Петроченко Никита, Коленько Александр, Сивов Георгий, Юртов Дмитрий; Группа - БИВ 214 , Команда № 6

## Предыстория

Плохое питание и болезни в детстве ограничивают рост человека. Как следствие, средний рост населения сильно коррелирует с уровнем жизни населения. Это делает изучение человеческого роста актуальным для историков, которые хотят понять историю условий жизни.

Поскольку эффект улучшения материального уровня жизни заключается в том, что люди становятся выше, человеческий рост используется в качестве косвенного показателя уровня жизни. Это особенно актуально для изучения условий жизни в периоды, для которых мало или вообще нет других данных, – то, что историки называют до-статистическим периодом.

Важно подчеркнуть, что рост не используется в качестве прямого показателя благополучия. Изменение роста в пределах данной популяции в значительной степени определяется генетическими факторами.<sup>1</sup>

История человеческого роста позволяет нам отслеживать прогресс в борьбе с недоеданием и болезнями и позволяет понять, кто и когда начал пользоваться преимуществами современных достижений.

## Оглавление

Отчет по Домашнему заданию №01 по предмету Проектный семинар “Python в науке о данных” .....	1
Тема работы: “Исследование зависимости роста человека по странам” .....	1
Предыстория .....	1
Нормальные формы базы данных .....	2
Графический анализ:.....	7

## Нормальные формы базы данных (1НФ) Первая нормальная форма:

ENTITY - название страны

YEAR - год сбора информации

CHANGE IN FEMALE HEIGHT (CM) FOR 100 YEARS - изменение в женском росте за 100 лет

CHANGE IN MALE HEIGHT (CM) FOR 100 YEARS - изменение в мужском росте за 100 лет

MEAN MALE HEIGHT FOR 100 YEARS- средний мужской рост

MEAN FEMALE HEIGHT FOR 100 YEARS - средний женский рост

CHANGE IN MALE HEIGHT(%) FOR 100 YEARS - изменение мужского роста в процентах

CHANGE IN FEMALE HEIGHT (%) FOR 100 YEARS- изменение женского роста в процентах

CALORIES FROM ANIMAL PROTEIN - количество белка, который люди в среднем потребили

HUMAN DEVELOPMENT INDEX (UNDP) - индекс уровня жизни (образование, здравоохранение и т.д.)

CONTINENT - Континент

CODE - Код страны

ENTITY	YEAR	CHANGE IN FEMALE HEIGHT (CM) FOR 100 YEARS	CHANGE IN MALE HEIGHT (CM) FOR 100 YEARS	MEAN MALE HEIGHT FOR 100 YEARS	MEAN FEMALE HEIGHT FOR 100 YEARS	CHANGE IN MALE HEIGHT(%) FOR 100 YEARS	CHANGE IN FEMALE HEIGHT (%) FOR 100 YEARS	CALORIES FROM ANIMAL PROTEIN	HUMAN DEVELOPMENT INDEX (UNDP)	CONTINENT	MORTALITY
Afghanistan	1996	5.8875762	4.0917654	165.2558607	155.0753231	2.5388815	3.94642075	64.04		Asia	14,56
Albania	1996	11.5737254	10.0637158	173.3888563	161.7698561	6.161767729	7.705741384	169.2	0.641	Europe	3,2600000000000002
Algeria	1996	7.2259589	9.683722	170.0718002	159.0896222	6.03768192	4.758188195	64.36	0.608	Africa	4,2025
American Samoa	1996	6.3239834	6.3294112	176.0451207	164.5213673	3.729419757	3.997527168			Oceania	
Andorra	1996	10.466057	12.67248	176.0601667	162.9048767	7.756080189	6.865742611			Europe	0,9400000000000001
Angola	1996	4.703628	5.1395795	167.3107273	157.3052103	3.169231747	3.082293073	40.52		Africa	21,97
Antigua and Barbuda	1996	10.0465608	10.9950508	172.7069903	160.6526419	6.799158327	6.670753748	203.64		North America	1,1959
Argentina	1996	8.5889525	12.2784569	174.6157252	159.1792058	7.563547809	5.703524838	257.2	0.738	South America	2,1931
Armenia	1996	8.4576857	10.1444336	171.9966462	158.0892096	6.267713884	5.652342153	68.48	0.612	Asia	3,7100000000000004
Australia	1996	10.366089	10.2573538	179.2046477	165.8581773	6.071333588	6.666634369	275.76	0.886	Oceania	0,68
Austria	1996	7.068036	9.3817916	177.4077047	164.6209454	5.583538531	4.486134865	261.84	0.82	Europe	0,63
Azerbaijan	1996	8.3191819	8.0795634	169.7494329	158.2524089	4.997569074	5.548591241	61.04	0.612	Asia	9,2018
Bahamas	1996	10.1188808	11.0186311	172.7466298	160.680362	6.813063408	6.720763318	183.8		North America	1,1977
Bahrain	1996	7.0309285	11.0318365	167.7434113	156.6916935	7.039579887	4.697910304		0.778	Asia	1,6600000000000001
Bangladesh	1996	9.3066267	5.2254168	163.8082283	150.7888441	3.295071358	6.577948007	23.04	0.433	Asia	10,840000000000002
Barbados	1996	13.217387	11.0649159	175.9161243	165.2770151	6.712062355	8.692239462	191.68	0.735	North America	1,4900000000000002
Belarus	1996	12.3789151	10.7062416	178.4384054	166.3465611	6.382938941	8.039945678	207.6	0.661	Europe	1,1955
Belgium	1996	10.069576	14.4983012	181.698803	165.4911355	8.671206751	6.478879785	252	0.857	Europe	0,71
Belize	1996	10.6951927	10.5999947	168.7317804	156.8758678	6.7032663	7.316420377	102.08	0.662	North America	2,7600000000000002
Benin	1996	2.9953795	4.5436044	167.0572709	156.1559783	2.795829113	1.955711536	30.52	0.377	Africa	15,200000000000001

Таблица 1. Исходное отношение

ETNITY\_CHANGEINMALE\_CHANGEINFEMALE\_CALORIES\_UNDP\_CONTINENT\_MORTALITY

Отношение в первой нормальной форме. Потенциальный ключ отношения - пара атрибутов {ENTITY, CHANGE IN HEIGHT}

(2НФ) Вторая нормальная форма

Наше исходное отношение уже находится во 2 Нормальной Форме

Зависимости :

ENTITY —> YEAR

ENTITY —> CHANGE IN FEMALE HEIGHT (CM) FOR 100 YEARS

ENTITY —> CHANGE IN MALE HEIGHT (CM) FOR 100 YEARS

ENTITY —> MEAN FEMALE HEIGHT FOR 100 YEARS

ENTITY —> MEAN MALE HEIGHT FOR 100 YEARS

ENTITY —> CHANGE IN MALE HEIGHT(%) FOR 100 YEARS

ENTITY —> CHANGE IN FEMALE HEIGHT(%) FOR 100 YEARS

ENTITY —> CALORIES FROM ANIMAL

ENTITY —> HUMAN DEVELOPMENT INDEX (UNDP)

ENTITY —> CODE

ENTITY —> CONTINENT

Далее, для того чтобы связать отношения, которые будут получены в приведении к ЗНФ, мы создали таблицу с отношением (СТРАНА,КОНТИНЕНТ\_КОД):

CODE	ENTITY	CONTINENT
AFG	Afghanistan	Asia
ALB	Albania	Europe
DZA	Algeria	Africa
ASM	American Samoa	Oceania
AND	Andorra	Europe
AGO	Angola	Africa
ATG	Antigua and Barbuda	North America
ARG	Argentina	South America
ARM	Armenia	Asia
AUS	Australia	Oceania
AUT	Austria	Europe
AZE	Azerbaijan	Asia
BHS	Bahamas	North America
BHR	Bahrain	Asia
BGD	Bangladesh	Asia
BRB	Barbados	North America
BLR	Belarus	Europe
BEL	Belgium	Europe
BLZ	Belize	North America
BEN	Benin	Africa
BMU	Bermuda	North America
BTN	Bhutan	Asia
BOL	Bolivia	South America
BIH	Bosnia and Herzegovina	Europe
BWA	Botswana	Africa
BRA	Brazil	South America
BRN	Brunei	Asia
BGR	Bulgaria	Europe

(ЗНФ) Третья нормальная форма

Тогда полученные отношения в третьей нормальной форме будут выглядеть так :

КОД\_ИНДЕКСЖИЗНИ\_СРЕДНИЙРОСТМУЖЧИНЫ

КОД\_СМЕРТНОСТЬ\_СРЕДНИЙРОСТМУЖЧИНЫ

КОД\_ИЗМЕНЕНИЕВРОСТЕМУЖЧИНЫ(%)\_ИЗМЕНЕНИЕВРОСТЕЖЕН  
ЩИНЫ(%)

КОД\_ИЗМЕНЕНИЕВРОСТЕМУЖЧИНЫ(СМ)\_ИЗМЕНЕНИЕВРОСТЕЖЕ  
НЩИНЫ(СМ)

КОД\_БЕЛОК\_СРЕДНИЙРОСТМУЖЧИНЫ

CODE	HUMAN DEVELOPMENT INDEX (UNDP)	MEAN MALE HEIGHT FOR 100 YEARS		
AFG		165.2558607		
ALB	0.641	173.3888563		
DZA	0.608	170.0718002		
ASM		176.0451207		
AND		176.0601667		
AGO		167.3107273		
ATG		172.7069903		
ARG	0.738	174.6157252		
ARM	0.612	171.9966462		
AUS	0.886	179.2046477		
AUT	0.82	177.4077047		
AZE	0.612	169.7494329		
BHS		172.7466298		
BHR	0.778	167.7434113		
BGD	0.433	163.8082283		
BRB	0.735	175.9161243		
BLR	0.661	178.4384054		
BEL	0.857	181.698803		
BLZ	0.662	168.7317804		
BEN	0.377	167.0572709		
BMU		172.6878702		
BTN		165.3058846		
BOL	0.578	166.8462997		
BIH		180.8691741		
BWA	0.571	171.6276505		
BRA	0.656	173.5512114		

Таблица 3 КОД\_ИНДЕКСЖИЗНИ\_СРЕДНИЙРОСТМУЖЧИНЫ

CODE	MORTALITY	MEAN MALE HEIGHT FOR 100 YEARS
AFG	14,56	165.2558607
ALB	3,2600000000000002	173.3888563
DZA	4,2025	170.0718002
ASM		176.0451207
AND	0,9400000000000001	176.0601667
AGO	21,97	167.3107273
ATG	1,1959	172.7069903
ARG	2,1931	174.6157252
ARM	3,7100000000000004	171.9966462
AUS	0,68	179.2046477
AUT	0,63	177.4077047
AZE	9,2018	169.7494329
BHS	1,1977	172.7466298
BHR	1,6600000000000001	167.7434113
BGD	10,840000000000002	163.8082283
BRB	1,4900000000000002	175.9161243
BLR	1,1955	178.4384054
BEL	0,71	181.698803
BLZ	2,7600000000000002	168.7317804
BEN	15,200000000000001	167.0572709
BMU		172.6878702
BTN	9,1956	165.3058846
BOL	9,3700000000000001	166.8462997
BIH	1,2022	180.8691741
BWA	6,6900000000000001	171.6276505
BRA	4,4799999999999995	173.5512114

Таблица 4 КОД\_СМЕРТНОСТЬ\_СРЕДНИЙРОСТМУЖЧИНЫ

CODE	CHANGE IN MALE HEIGHT(%) FOR 100 YEARS	CHANGE IN FEMALE HEIGHT (%) FOR 100 YEARS
AFG	2.5388815	3.94642075
ALB	6.161767729	7.705741384
DZA	6.03768192	4.758188195
ASM	3.729419757	3.997527168
AND	7.756080189	6.865742611
AGO	3.169231747	3.082293073
ATG	6.799158327	6.670753748
ARG	7.563547809	5.703524838
ARM	6.267713884	5.652342153
AUS	6.071333588	6.666634369
AUT	5.583538531	4.486134865
AZE	4.997569074	5.548591241
BHS	6.813063408	6.720763318
BHR	7.039579887	4.697910304
BGD	3.295071358	6.577948007
BRB	6.712062355	8.692239462
BLR	6.382938941	8.039945678
BEL	8.671206751	6.478879785
BLZ	6.7032663	7.316420377
BEN	2.795829113	1.955711536
BMU	6.763113416	6.755970258
BTN	2.504867163	4.530942747
BOL	6.735351769	7.467855104
BIH	7.389369557	7.80250508
BWA	3.274247957	3.427269489
BRA	6.333814394	7.122724804
BRN	5.918493461	5.653879315
BGR	7.078215606	8.282707059

Таблица 5 КОД\_ИЗМЕНЕНИЕВРОСТЕМУЖЧИНЫ(%)\_ИЗМЕНЕНИЕВРОСТЕЖЕНЩИНЫ(%)

CODE	CHANGE IN FEMALE HEIGHT (CM) FOR 100 YEARS	CHANGE IN MALE HEIGHT (CM) FOR 100 YEARS
AFG	5.8875762	4.0917654
ALB	11.5737254	10.0637158
DZA	7.2259589	9.683722
ASM	6.3239834	6.3294112
AND	10.466057	12.67248
AGO	4.703628	5.1395795
ATG	10.0465608	10.9950508
ARG	8.5889525	12.2784569
ARM	8.4576857	10.1444336
AUS	10.366089	10.2573538
AUT	7.068036	9.3817916
AZE	8.3191819	8.0795634
BHS	10.1188808	11.0186311
BHR	7.0309285	11.0318365
BGD	9.3066267	5.2254168
BRB	13.217387	11.0649159
BLR	12.3789151	10.7062416
BEL	10.069576	14.4983012
BLZ	10.6951927	10.5999947
BEN	2.9953795	4.5436044
BMU	10.1689983	10.9392431
BTN	6.6591408	4.0395085
BOL	10.6935803	10.5285503
BIH	12.0037311	12.4454513
BWA	5.3476914	5.4413515
BRA	10.6960196	10.3376444
BRN	8.2400987	9.2203452
BGR	12.6054448	11.7822558

Таблица 6 КОД\_ИЗМЕНЕНИЕВРОСТЕМУЖЧИНЫ(СМ)\_ИЗМЕНЕНИЕВРОСТЕЖЕНЩИНЫ(СМ)

CODE	CALORIES FROM ANIMAL PROTEIN (FAO (2017))	MEAN MALE HEIGHT FOR 100 YEARS
AFG	64.04	165.2558607
ALB	169.2	173.3888563
DZA	64.36	170.0718002
ASM		176.0451207
AND		176.0601667
AGO	40.52	167.3107273
ATG	203.64	172.7069903
ARG	257.2	174.6157252
ARM	68.48	171.9966462
AUS	275.76	179.2046477
AUT	261.84	177.4077047
AZE	61.04	169.7494329
BHS	183.8	172.7466298
BHR		167.7434113
BGD	23.04	163.8082283
BRB	191.68	175.9161243
BLR	207.6	178.4384054
BEL	252	181.698803
BLZ	102.08	168.7317804
BEN	30.52	167.0572709
BMU	224.76	172.6878702
BTN		165.3058846
BOL	89.32	166.8462997
BIH	82.96	180.8691741
BWA	106.72	171.6276505
BRA	155.56	173.5512114
BRN	205.48	165.0090651
BGR	155.32	178.2402511

Таблица 7 КОД\_БЕЛОК\_СРЕДНИЙРОСТМУЖЧИНЫ

Code	mean male height	mean female height
AFG	165.2558607	155.0753231
ALB	173.3888563	161.7698561
DZA	170.0718002	159.0896222
ASM	176.0451207	164.5213673
AND	176.0601667	162.9048767
AGO	167.3107273	157.3052103
ATG	172.7069903	160.6526419
ARG	174.6157252	159.1792058
ARM	171.9966462	158.0892096
AUS	179.2046477	165.8581773
AUT	177.4077047	164.6209454
AZE	169.7494329	158.2524089
BHS	172.7466298	160.680362
BHR	167.7434113	156.6916935
BGD	163.8082283	150.7888441
BRB	175.9161243	165.2770151
BLR	178.4384054	166.3465611
BEL	181.698803	165.4911355
BLZ	168.7317804	156.8758678
BEN	167.0572709	156.1559783
BMU	172.6878702	160.6876938
BTN	165.3058846	153.6294552
BOL	166.8462997	153.8883819
BIH	180.8691741	165.8483102
BWA	171.6276505	161.3812749
BRA	173.5512114	160.8635453
BRN	165.0090651	153.9824862
BGR	178.2402511	164.7953594
BFA	169.3288328	160.1857488
BDI	166.6352493	154.0233342

Таблица 8 КОД\_СРРОСТМУЖЧИН\_СРРОСТЖЕНЩИН

В результате, получена структура из изначальной базы данных ,которая представляет собой 5 взаимосвязанных отношений , на основе которых уже можно делать графический анализ .

### Графический анализ:

1. Точечный график, имеющий заголовок «Процент возрастания среднего роста мужчин и женщин за 100 лет» (Рисунок 2) по странам по данным таблицы «Изменение в росте.xlsx» (Таблица 4). Точка показывает на процент возрастания в данной стране.

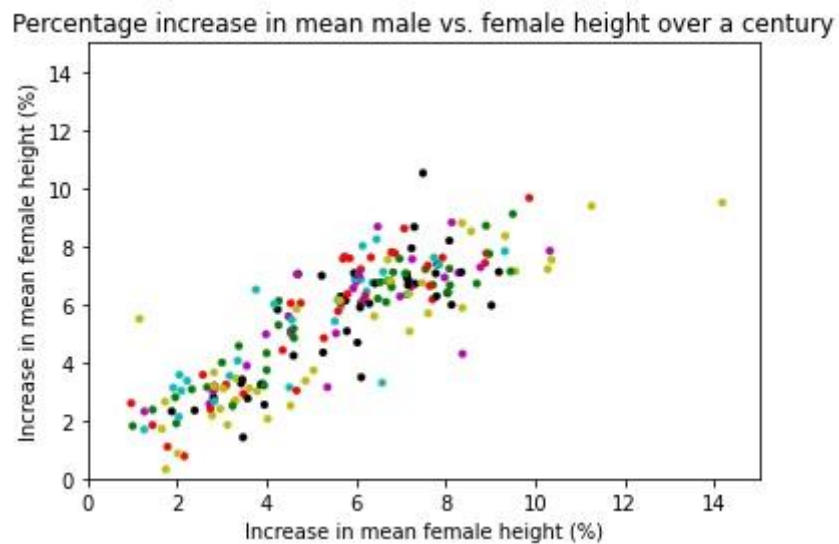


Рисунок 2

2. Столбчатая диаграмма, имеющая заголовок «Средний рост на различных континентах» (Рисунок 3) по континентам и по данным таблицы «Изменение в росте мужчин и женщин.xlsx» (Таблица 6). Верхняя граница графика - средний рост по континенту.

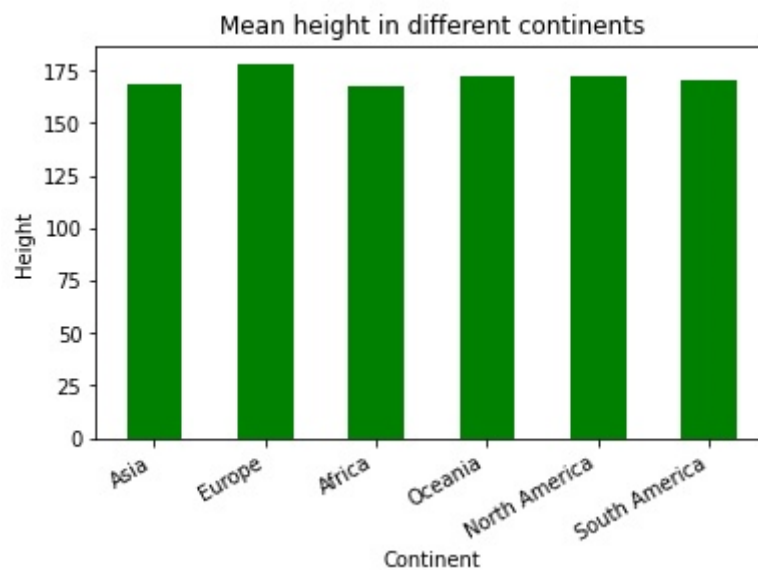


Рисунок 3



3. Круговая диаграмма, имеющая заголовок «Процент представителей из различных континентов» (Рисунок 4) по данным таблицы «Рост Мужчин и Женщин.xlsx» (Таблица 8). Каждый «кусочек» - количество представителей, населяющих континенты, данные о которых представлены в нашей таблице.

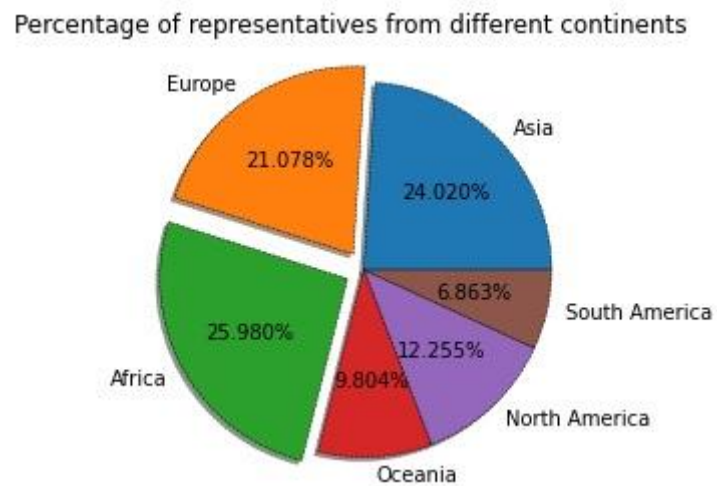


Рисунок 4