

PONTIFÍCIA UNIVERSIDADE CATÓLICA DE MINAS GERAIS
NÚCLEO DE EDUCAÇÃO A DISTÂNCIA
Pós-graduação *Lato Sensu* em Ciência de Dados e Big Data

Manuel Antonio Rodrigues Leal

Dados do Vigitel e os números da obesidade no Brasil

Belo Horizonte
2020

Manuel Antonio Rodrigues Leal

Dados do Vigitel e os números da obesidade no Brasil

Trabalho de Conclusão de Curso apresentado
ao Curso de Especialização em Ciência de
Dados e Big Data como requisito parcial à
obtenção do título de especialista.

Belo Horizonte
2020

SUMÁRIO

1. Introdução.....	4
1.1. Contextualização.....	4
1.2. O problema proposto.....	5
2. Coleta de Dados.....	6
3. Processamento/Tratamento de Dados.....	7
3.1. Quantidade de registros importados.....	7
3.2. Registros por ano e cidade.....	8
4. Análise e Exploração dos Dados.....	10
4.1 Evolução da obesidade.....	10
4.1 Evolução da obesidade por sexo.....	12
4.2 Distribuição da obesidade por local.....	15
4.3 Evolução anual de doenças.....	17
4.4 Relação pressão alta e obesidade (2019).....	18
4.5 Relação diabetes e obesidade (2019).....	19
5. Apresentação dos Resultados.....	21
5.1 Workflow.....	21
5.2 Evolução das doenças.....	21
6. Links.....	22
6.1 Vídeo de apresentação.....	22
6.2 Repositório do Projeto.....	22
REFERÊNCIAS.....	23

1. Introdução

1.1. Contextualização

Para saber os hábitos, comportamentos e situação de saúde dos brasileiros, a Secretaria de Vigilância em Saúde (SVS), do **Ministério da Saúde**, realiza anualmente a Pesquisa de **Vigilância de Fatores de Risco e Proteção para Doenças crônicas por Inquérito Telefônico (Vigitel) (1)**.

Conhecer a situação de saúde da população é o primeiro passo para planejar ações e programas que reduzam a ocorrência e a gravidade destas doenças, melhorando assim a saúde da população **(2)**.

Os participantes são perguntados sobre hábito de fumar, consumo alimentar e de bebidas alcoólicas, prática de atividade física, presença de diabetes e hipertensão. Além disso, são realizadas outras perguntas que não são diretamente sobre saúde, mas são muito importantes para serem relacionadas com a situação da saúde da população.

Neste trabalho usaremos dados do **Vigitel** para analisar a evolução de doenças crônicas, e a relação entre elas no Brasil, principalmente a **obesidade**.

Obesidade

O **índice de massa corporal (IMC) (3)** é uma medida internacional usada para calcular se uma pessoa está no peso ideal. Desenvolvido no fim do século XIX, trata-se de um método fácil e rápido para a avaliação do nível de gordura de cada pessoa, sendo, por isso, um preditor internacional de obesidade adotado pela **Organização Mundial da Saúde (OMS)**.

$$\text{IMC} = \frac{\text{massa}}{(\text{altura} \times \text{altura})} \text{ kg/m}^2$$

Cálculo exemplo:

Por exemplo, uma pessoa que tenha 1,75 metros de altura e 83 kg de massa **(4)** possui o seguinte **IMC**:

$$\text{IMC} = 83 / ((1,75) * (1,75)) = \mathbf{27,1 \text{ kg/m}^2}$$

Classificações (faixas) do IMC

IMC	Classificações
Menor do que 18,5	Abaixo do peso normal
18,5 - 24,9	Peso normal
25,0 - 29,9	Excesso de peso
30,0 - 34,9	Obesidade classe I
35,0 - 39,9	Obesidade classe II
Maior ou igual a 40,0	Obesidade classe III

Classificação segundo a OMS a partir do IMC

Para nossa análise, iremos considerar como **Obesidade** o **IMC** a partir do valor 30 (inclusive). O **IMC** de 25 (inclusive) a 30 (exclusive) está acima do normal, mas é classificado como **sobrepeso** ou pré-obeso.

Obesidade → IMC ≥ 30 kg/m²

1.2. O problema proposto

Analisar a evolução e a relação de doenças, principalmente a **obesidade**.

(Why?) Conhecer a situação de saúde da população brasileira.

(Who?) Dados do **Vigitel** do Ministério da Saúde do governo brasileiro.

(What?) Para mapear as relações de doenças x evolução x hábitos.

(Where?) Amostras coletadas nas capitais brasileiras e no Distrito Federal.

(When?) Coletas anuais por telefone desde 2006.

2. Coleta de Dados

Os dados são de consulta pública e podem ser obtidos em formato de planilha (.xls) ou texto (.csv) pela internet (5). A descrição das variáveis disponíveis para tabulação estão em anexo neste documento (6).

Para este trabalho os campos necessários foram importados para uma tabela (**VIGITEL**) de um arquivo *SQLite* local (**DB_VIGITEL.db**). Um módulo em *Python* (**vigitel.py**) foi criado para realizar importações de planilhas anuais, de 2009 a 2019 ("**Vigitel-2009-peso-rake.xls**" ... "**Vigitel-2019-peso-rake.xls**"). Embora existam coletas desde 2006 ("**Vigitel-2006-peso-rake.xls**") (5), anos anteriores a 2009 não serão usados neste trabalho.

Tabela **VIGITEL** de **DB_VIGITEL.db** (*SQLite*)

CAMPO	TIPO	ORIGEM	DESCRICAO
ID	INTEGER	-----	primary key
ANO	INTEGER	ano	ano
CIDADE	INTEGER	cidade	cidade
IDADE	INTEGER	q6	idade (anos)
SEXO	INTEGER	q7	sexo
CIVIL	INTEGER	civil	estado conjugal atual
ESTUDO_ANOS	INTEGER	q8_anos	anos de estudo
PESO	REAL	q9	peso (kg)
ALTURA	INTEGER	q11	altura (cm)
EXERCICIO_FISICO	INTEGER	q42	exercício físico
EXERCICIO_FREQ	INTEGER	q45	freq exercicio
FUMANTE	INTEGER	q60	fumante
COR	INTEGER	q69	cor
PRESSAO_ALTA	INTEGER	q75	pressão alta
DIABETES	INTEGER	q76 ou q76a	diabetes
IMC	REAL	imc	imc

3. Processamento/Tratamento de Dados

Os dados (amostras) foram importados de 11 arquivos (planilhas de *Excel*), um arquivo de cada ano (de 2009 a 2019), para uma única tabela local (**VIGITEL**). Como são dados “prontos” para análise, e disponíveis para uso público, é de se esperar que não tragam muitos *outliers* ou registros duplicados, mesmo assim os mesmos serão tratados nas respectivas análises. Alguns valores serão excluídos no processo de consulta, por exemplo valor 777 (“não sabe”) para a variável **peso**. Valores nulos (não coletados em algumas ligações telefônicas) também serão tratados de acordo com a consulta feita.

Todas as variáveis foram importadas com seus valores originais, por isso foram criados dicionários para algumas delas (cidade, sexo, civil, peso, altura, exercicio_fisico, exercicio_freq, fumante, cor, pressao_alta e diabetes) em **vigitel.py**.

As consultas e análises foram trabalhadas num arquivo *jupyter notebook* (**Analise_Vigitel.ipynb**), com pivotamento de *DataFrames* quando necessário. As consultas a tabela **VIGITEL** (banco de dados SQLite) são feitas por *script* de linguagem SQL (estão neste documento também). Foi usada uma planilha para ajudar em alguns cálculos pontuais e melhorar a aparência de tabelas e gráficos (**calculo_aux.xlsx**).

3.1 Quantidade de registros importados

```
SELECT ANO, COUNT(*) AS REGISTROS FROM VIGITEL GROUP BY ANO
```

ANO	ORIGEM	REGISTROS	
2009	Vigitel-2009-peso-rake.xls	54.367	9,6%
2010	Vigitel-2010-peso-rake.xls	54.339	9,6%
2011	Vigitel-2011-peso-rake.xls	54.144	9,5%
2012	Vigitel-2012-peso-rake.xls	45.448	8,0%
2013	Vigitel-2013-peso-rake.xls	52.929	9,3%
2014	Vigitel-2014-peso-rake.xls	40.853	7,2%
2015	Vigitel-2015-peso-rake.xls	54.174	9,5%
2016	Vigitel-2016-peso-rake.xls	53.210	9,4%
2017	Vigitel-2017-peso-rake.xls	53.034	9,3%
2018	Vigitel-2018-peso-rake.xls	52.395	9,2%
2019	Vigitel-2019-peso-rake.xls	52.443	9,2%
TOTAL		567.336	100,0%

Há uma pequena diferença nas quantidades coletas nos anos de 2012 e 2014, com relação aos demais anos. Não parece que impactarão nas análises a serem feitas, mas é prudente ter esse fato em mente.

3.2 Registros por ano e cidade

```
SELECT ANO, CIDADE, COUNT(*) AS REGISTROS FROM VIGITEL GROUP BY ANO, CIDADE
```

	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	TOTAL	
aracaju	2.014	2.006	2.003	1.677	1.942	1.503	2.005	2.015	2.025	2.001	2.053	21.244	3,7%
belem	2.012	2.009	2.042	1.696	1.955	1.512	2.003	1.998	1.920	2.000	2.059	21.206	3,7%
belo horizonte	2.011	2.007	2.006	1.811	1.956	1.519	2.007	2.004	2.008	2.072	2.064	21.465	3,8%
boa vista	2.010	2.019	2.018	1.692	1.953	1.500	2.026	1.853	1.674	1.200	1.017	18.962	3,3%
campo grande	2.013	2.010	2.000	1.666	1.949	1.501	2.007	2.011	2.032	2.000	2.086	21.275	3,7%
cuiaba	2.012	2.010	2.001	1.666	1.964	1.509	2.006	2.040	2.026	2.010	2.064	21.308	3,8%
curitiba	2.014	2.010	2.000	1.668	1.951	1.504	2.002	2.011	2.063	2.058	2.058	21.339	3,8%
florianopolis	2.011	2.007	2.000	1.669	1.956	1.507	2.005	1.886	1.970	2.005	2.080	21.096	3,7%
fortaleza	2.011	2.011	2.002	1.690	1.977	1.518	2.001	1.944	2.027	2.075	2.051	21.307	3,8%
goiania	2.011	2.012	2.000	1.710	1.979	1.538	2.003	2.012	2.048	2.047	2.050	21.410	3,8%
joao pessoa	2.009	2.011	2.002	1.667	1.953	1.517	2.000	1.919	2.063	2.002	2.064	21.207	3,7%
macapa	2.010	2.004	2.000	1.650	1.949	1.500	2.011	1.880	1.439	1.333	973	18.749	3,3%
maceio	2.012	2.007	2.001	1.668	1.978	1.518	2.005	2.017	2.056	2.051	2.064	21.377	3,8%
manaus	2.021	2.098	2.015	1.629	1.959	1.502	2.006	2.014	1.646	1.601	1.953	20.444	3,6%
natal	2.024	2.006	2.000	1.698	1.956	1.509	2.017	2.009	2.018	2.048	2.076	21.361	3,8%
palmas	2.012	2.009	2.002	1.692	1.960	1.514	2.002	2.034	2.070	2.018	1.470	20.783	3,7%
porto alegre	2.010	2.005	2.016	1.665	1.949	1.506	2.013	2.035	2.067	2.040	2.058	21.364	3,8%
porto velho	2.025	2.008	2.000	1.668	1.954	1.513	2.004	1.897	1.846	2.048	1.926	20.889	3,7%
recife	2.010	2.005	2.016	1.671	1.951	1.505	2.010	2.031	2.053	2.066	2.053	21.371	3,8%
rio branco	2.023	2.016	2.002	1.662	1.971	1.517	2.011	1.806	1.831	1.461	1.812	20.112	3,5%
rio de janeiro	2.011	2.007	2.004	1.690	1.980	1.537	2.012	1.934	2.031	2.062	2.059	21.327	3,8%
salvador	2.020	2.009	2.001	1.685	1.960	1.533	2.002	1.922	2.048	2.030	2.056	21.266	3,7%
sao luis	2.011	2.021	2.009	1.675	1.942	1.512	2.005	1.934	2.051	2.004	2.065	21.229	3,7%
sao paulo	2.010	2.008	2.001	1.737	1.999	1.535	2.002	2.034	2.020	2.052	2.052	21.450	3,8%
teresina	2.011	2.010	2.001	1.667	1.954	1.500	2.001	2.001	2.041	2.017	2.075	21.278	3,8%
vitoria	2.019	2.010	2.000	1.679	1.966	1.511	2.001	2.003	2.055	2.002	2.051	21.297	3,8%
distrito federal	2.010	2.004	2.002	1.700	1.966	1.513	2.007	1.966	1.906	2.092	2.054	21.220	3,7%
	54.367	54.339	54.144	45.448	52.929	40.853	54.174	53.210	53.034	52.395	52.443	567.336	100,0%
	9,6%	9,6%	9,5%	8,0%	9,3%	7,2%	9,5%	9,4%	9,3%	9,2%	9,2%	100,0%	

A distribuição da quantidade de registros parece equilibrada para uma análise comparativa, levando em consideração tempo (**anos**) e espaço (**cidades**). Acrescentando mais uma variável (**sexo**) para reforçar a ideia de consistência da consolidação das amostras:

```
SELECT ANO, CIDADE, SEXO, COUNT(*) AS REGISTROS FROM VIGITEL GROUP BY ANO, CIDADE, SEXO
```


		2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
aracaju	masculino	799	744	762	609	762	536	745	765	706	668	702
	feminino	1215	1262	1241	1068	1180	967	1260	1250	1319	1333	1351
belem	masculino	776	751	804	653	732	578	809	735	741	706	731
	feminino	1236	1258	1238	1043	1223	934	1194	1263	1179	1294	1328
belo horizonte	masculino	817	789	836	689	757	527	779	747	816	772	766
	feminino	1194	1218	1170	1122	1199	992	1228	1257	1192	1300	1298
boa vista	masculino	812	825	815	657	790	652	808	794	680	468	389
	feminino	1198	1194	1203	1035	1163	848	1218	1059	994	732	628
campo grande	masculino	771	788	785	658	727	546	785	755	714	725	731
	feminino	1242	1222	1215	1008	1222	955	1222	1256	1318	1275	1355
cuiaba	masculino	806	809	833	649	747	575	659	792	698	797	717
	feminino	1206	1201	1168	1017	1217	934	1347	1248	1328	1213	1347
curitiba	masculino	796	761	812	644	730	557	830	737	658	736	737
	feminino	1218	1249	1188	1024	1221	947	1172	1274	1405	1322	1321
florianopolis	masculino	801	757	795	676	742	573	710	677	678	762	790
	feminino	1210	1250	1205	993	1214	934	1295	1209	1292	1243	1290
fortaleza	masculino	738	734	776	656	785	557	739	687	725	708	702
	feminino	1273	1277	1226	1034	1192	961	1262	1257	1302	1367	1349
goiania	masculino	808	785	760	645	746	583	638	755	704	717	720
	feminino	1203	1227	1240	1065	1233	955	1365	1257	1344	1330	1330
joao pessoa	masculino	733	697	790	590	702	579	708	641	746	638	636
	feminino	1276	1314	1212	1077	1251	938	1292	1278	1317	1364	1428
macapa	masculino	796	816	820	650	765	624	792	734	559	496	365
	feminino	1214	1188	1180	1000	1184	876	1219	1146	880	837	608
maceio	masculino	783	730	746	621	750	533	765	740	693	704	634
	feminino	1229	1277	1255	1047	1228	985	1240	1277	1363	1347	1430
manaus	masculino	787	822	784	623	822	599	759	835	623	575	707
	feminino	1234	1276	1231	1006	1137	903	1247	1179	1023	1026	1246
natal	masculino	771	732	779	625	705	552	748	730	742	739	711
	feminino	1253	1274	1221	1073	1251	957	1269	1279	1276	1309	1365
palmas	masculino	865	841	872	726	831	654	877	914	873	872	566
	feminino	1147	1168	1130	966	1129	860	1125	1120	1197	1146	904
porto alegre	masculino	771	701	712	573	700	536	793	752	709	668	710
	feminino	1239	1304	1304	1092	1249	970	1220	1283	1358	1372	1348
porto velho	masculino	879	843	893	682	861	614	797	778	766	832	744
	feminino	1146	1165	1107	986	1093	899	1207	1119	1080	1216	1182
recife	masculino	756	762	724	606	676	532	717	716	702	712	690
	feminino	1254	1243	1292	1065	1275	973	1293	1315	1351	1354	1363
rio branco	masculino	788	754	783	638	788	589	712	770	699	524	631
	feminino	1235	1262	1219	1024	1183	928	1299	1036	1132	937	1181
rio de janeiro	masculino	758	744	746	642	733	562	727	733	825	767	730
	feminino	1253	1263	1258	1048	1247	975	1285	1201	1206	1295	1329
salvador	masculino	789	753	810	607	728	592	712	704	738	729	641
	feminino	1231	1256	1191	1078	1232	941	1290	1218	1310	1301	1415
sao luis	masculino	794	761	782	650	713	595	719	752	745	732	707
	feminino	1217	1260	1227	1025	1229	917	1286	1182	1306	1272	1358
sao paulo	masculino	757	748	829	679	775	588	805	783	748	766	708
	feminino	1253	1260	1172	1058	1224	947	1197	1251	1272	1286	1344
teresina	masculino	752	767	780	635	753	562	769	782	760	754	733
	feminino	1259	1243	1221	1032	1201	938	1232	1219	1281	1263	1342
vitoria	masculino	812	749	784	633	703	570	683	669	728	690	700
	feminino	1207	1261	1216	1046	1263	941	1318	1334	1327	1312	1351
distrito federal	masculino	832	801	814	673	753	556	783	781	728	782	756
	feminino	1178	1203	1188	1027	1213	957	1224	1185	1178	1310	1298

4. Análise e Exploração dos Dados

4.1 Evolução da obesidade

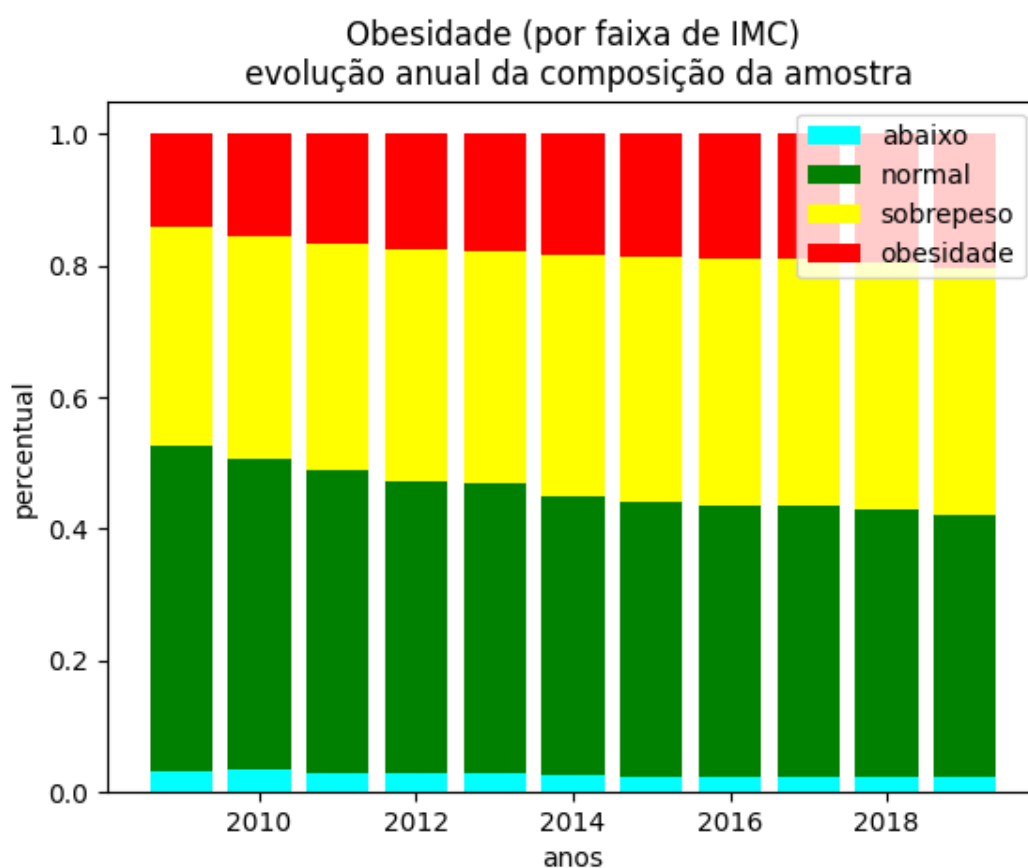
As amostras do **VIGITEL** parecem ser bem consistentes e equilibradas para começarmos a tentar responder uma pergunta: será que o índice de **obesidade** aumentou, como o senso comum sugere?

Podemos agrupar por **ano** e por faixa de **IMC (obesidade)** a tabela **VIGITEL**. Vamos chamar de valores **VALIDOS**, valores de **IMC** não zerados, não nulos e abaixo de 100 kg/m².

```
SELECT ANO
      , SUM(VALIDOS)      AS VALIDOS
      , SUM(ABAIXO)      AS ABAIXO
      , SUM(NORMAL)      AS NORMAL
      , SUM(SOBREPESO)    AS SOBREPESO
      , SUM(OBESIDADE)    AS OBESIDADE
FROM (SELECT ANO
      , (CASE WHEN (IMC > 0 AND IMC < 100) THEN 1 ELSE 0 END) AS VALIDOS
      , (CASE WHEN (IMC > 0 AND IMC < 18.5) THEN 1 ELSE 0 END) AS ABAIXO
      , (CASE WHEN (IMC >= 18.5 AND IMC < 25) THEN 1 ELSE 0 END) AS NORMAL
      , (CASE WHEN (IMC >= 25 AND IMC < 30) THEN 1 ELSE 0 END) AS SOBREPESO
      , (CASE WHEN (IMC >= 30 AND IMC < 100) THEN 1 ELSE 0 END) AS OBESIDADE
FROM VIGITEL) V
GROUP BY ANO
ORDER BY ANO
```

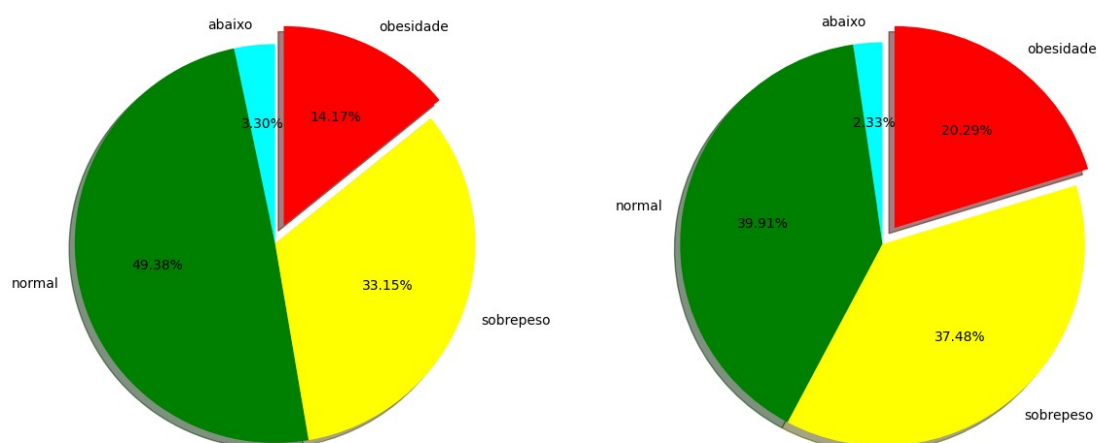
ANO	VALIDOS	ABAIXO	NORMAL	SOBREPESO	OBESIDADE
2009	49.972	1.647	24.677	16.565	7.083
2010	49.682	1.668	23.468	16.777	7.769
2011	49.411	1.466	22.757	16.982	8.206
2012	41.527	1.166	18.476	14.591	7.294
2013	48.265	1.349	21.360	16.974	8.582
2014	37.112	948	15.769	13.594	6.801
2015	48.653	1.115	20.339	18.124	9.075
2016	48.911	1.120	20.147	18.369	9.275
2017	48.305	1.149	19.875	18.060	9.221
2018	47.708	1.114	19.321	17.892	9.381
2019	47.119	1.096	18.804	17.658	9.561

O índice de valores **validos**, na amostra, é de **91,1%** do total de registros (516.665 de 567.336).



O gráfico mostra uma diminuição da parcela que possui peso **normal** (verde) e aumento nas faixas de **sobrepeso** (amarelo) e **obesidade** (vermelho) no período analisado (uma década). Podemos fazer um confronto entre o primeiro e o último ano:

Composição da amostra: 2009 x 2019



Alterações ocorridas:

- **abaixo** (ciano), diminuição de 3,30% para 2,33%
- **normal** (verde), diminuição de 49,38% para 39,91%
- **sobrepeso** (amarelo), aumento de 33,15% para 37,48%
- **obesidade** (vermelho), aumento de 14,17% para 20,29%

4.1 Evolução da obesidade por sexo

Analizando a diferença, entre os sexos, do impacto **obesidade**.

```

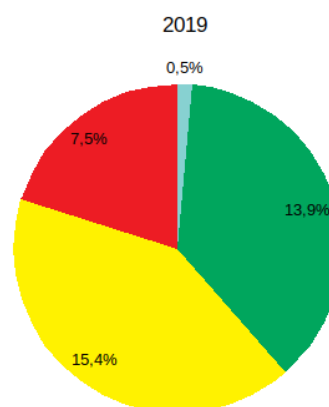
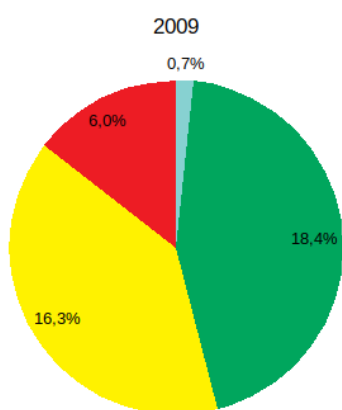
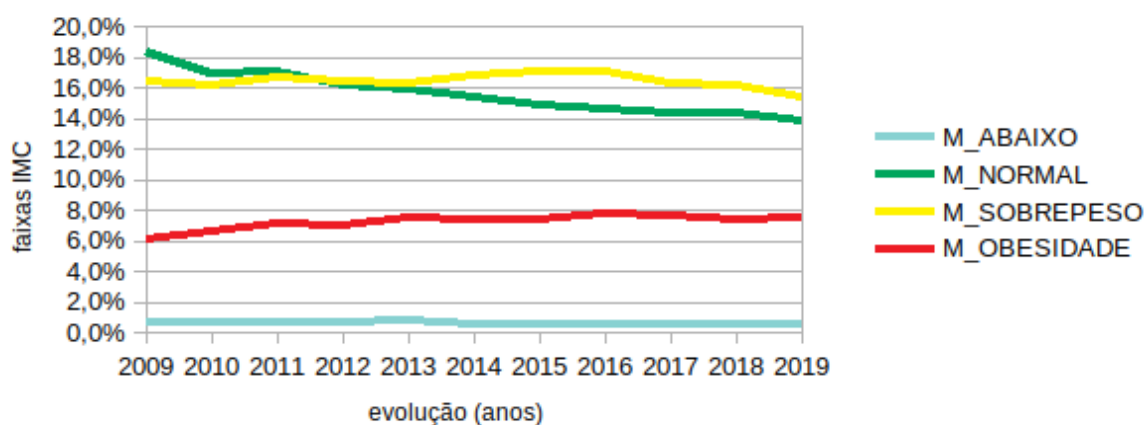
SELECT ANO
      , SUM(M_ABAIXO)      AS M_ABAIXO
      , SUM(M_NORMAL)      AS M_NORMAL
      , SUM(M_SOBREPESO)   AS M_SOBREPESO
      , SUM(M_OBESIDADE)   AS M_OBESIDADE
      , SUM(F_ABAIXO)      AS F_ABAIXO
      , SUM(F_NORMAL)      AS F_NORMAL
      , SUM(F_SOBREPESO)   AS F_SOBREPESO
      , SUM(F_OBESIDADE)   AS F_OBESIDADE
FROM(
SELECT ANO
      , (CASE WHEN (IMC > 0 AND IMC < 100 AND
                    SEXO=1) THEN 1 ELSE 0 END) AS M_VALIDOS
      , (CASE WHEN (IMC > 0 AND IMC < 18.5 AND
                    SEXO=1) THEN 1 ELSE 0 END) AS M_ABAIXO
      , (CASE WHEN (IMC >= 18.5 AND IMC < 25 AND
                    SEXO=1) THEN 1 ELSE 0 END) AS M_NORMAL
      , (CASE WHEN (IMC >= 25 AND IMC < 30 AND
                    SEXO=1) THEN 1 ELSE 0 END) AS M_SOBREPESO
      , (CASE WHEN (IMC >= 30 AND IMC < 100 AND
                    SEXO=1) THEN 1 ELSE 0 END) AS M_OBESIDADE
      , (CASE WHEN (IMC > 0 AND IMC < 100 AND
                    SEXO=2) THEN 1 ELSE 0 END) AS F_VALIDOS
      , (CASE WHEN (IMC > 0 AND IMC < 18.5 AND
                    SEXO=2) THEN 1 ELSE 0 END) AS F_ABAIXO
      , (CASE WHEN (IMC >= 18.5 AND IMC < 25 AND
                    SEXO=2) THEN 1 ELSE 0 END) AS F_NORMAL
      , (CASE WHEN (IMC >= 25 AND IMC < 30 AND
                    SEXO=2) THEN 1 ELSE 0 END) AS F_SOBREPESO
      , (CASE WHEN (IMC >= 30 AND IMC < 100 AND
                    SEXO=2) THEN 1 ELSE 0 END) AS F_OBESIDADE
FROM VIGITEL) V
GROUP BY ANO

```

Evolução da composição da **obesidade** na população **masculina**

ANO	M_VALIDOS	M_ABAIXO	M_NORMAL	M_SOBREPESO	M_OBESIDADE
2009	20.693	340	9.171	8.160	3.022
2010	19.996	338	8.382	8.012	3.264
2011	20.580	360	8.442	8.251	3.527
2012	16.683	284	6.677	6.795	2.927
2013	19.489	364	7.650	7.824	3.651
2014	14.906	226	5.724	6.237	2.719
2015	19.388	264	7.234	8.293	3.597
2016	19.546	271	7.157	8.336	3.782
2017	18.772	293	6.941	7.846	3.692
2018	18.314	270	6.828	7.672	3.544
2019	17.579	253	6.528	7.260	3.538

Evolução da composição da Obesidade

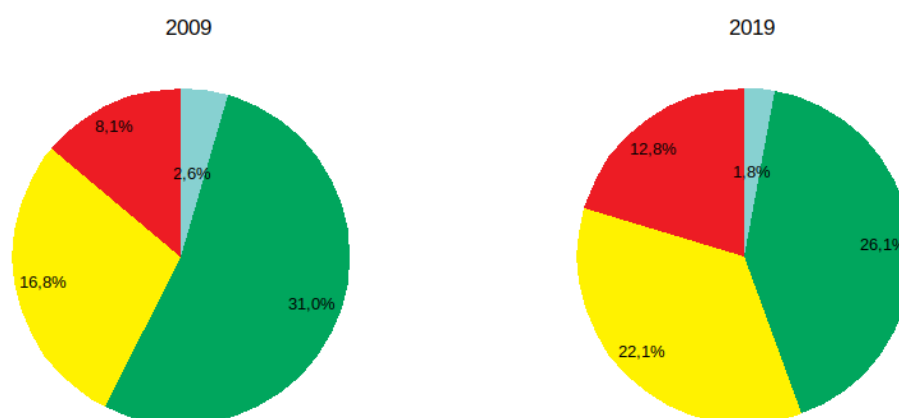
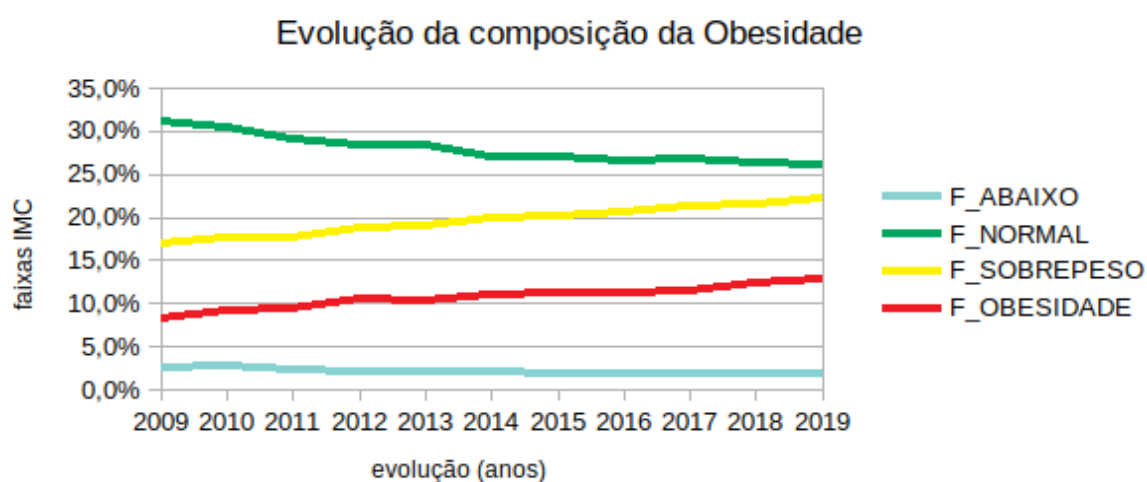


Alterações ocorridas:

- **abaixo** (ciano), diminuição de 0,7% para 0,5%
- **normal** (verde), diminuição de 18,4% para 13,9%
- **sobrepeso** (amarelo), diminuição de 16,3% para 15,4%
- **obesidade** (vermelho), aumento de 6,0% para 7,5%

Evolução da composição da **obesidade** na população **feminina**

ANO	M_VALIDOS	F_ABAIXO	F_NORMAL	F_SOBREPESO	F_OBESIDADE
2009	29.279	1.307	15.506	8.405	4.061
2010	29.686	1.330	15.086	8.765	4.505
2011	28.831	1.106	14.315	8.731	4.679
2012	24.844	882	11.799	7.796	4.367
2013	28.776	985	13.710	9.150	4.931
2014	22.206	722	10.045	7.357	4.082
2015	29.265	851	13.105	9.831	5.478
2016	29.365	849	12.990	10.033	5.493
2017	29.533	856	12.934	10.214	5.529
2018	29.394	844	12.493	10.220	5.837
2019	29.540	843	12.276	10.398	6.023



Alterações ocorridas:

- **abaixo** (ciano), diminuição de 2,6% para 1,8%
- **normal** (verde), diminuição de 31,0% para 26,1%
- **sobrepeso** (amarelo), aumento de 16,8% para 22,1%
- **obesidade** (vermelho), aumento de 8,1% para 12,8%

Os percentuais dos **gráficos pizza** são pesos das faixas de IMC do sexo analisado em relação ao total válido (masculino + feminino), ajudam a entender a alteração ocorrida entre 2009 e 2019.

Em 2009 a **obesidade** era maior nas mulheres 8,1%, contra 6,0% nos homens. Em 2019 a **obesidade aumentou** em ambos os sexos, mas o maior aumento foi nas mulheres 12,8%, contra 7,5% nos homens.

Vale destacar que os índices de **abaixo** e **normal** caíram em ambos os sexos, mas o índice de **sobrepeso** se comportou de forma diferente, aumentou no sexo feminino, de 16,8% para 22,1%, e diminuiu no sexo masculino 16,3% para 15,4%.

4.2 Distribuição da obesidade por local

Algum local conseguiu diminuir o índice de **obesidade**? Quais as cidades de maior e menor índice?

Evolução anual por local

	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
aracaju	15,5%	16,1%	15,5%	16,1%	18,7%	16,9%	17,6%	20,5%	19,3%	19,1%	20,8%
belem	13,1%	15,6%	15,1%	17,1%	16,5%	18,4%	18,7%	20,1%	18,9%	18,8%	20,1%
belo horizonte	12,5%	12,5%	14,4%	15,7%	15,5%	16,5%	18,1%	17,2%	16,2%	18,7%	19,8%
boa vista	14,6%	14,9%	16,5%	18,2%	18,3%	19,3%	19,1%	17,2%	19,5%	19,7%	22,7%
campo grande	16,6%	18,0%	18,5%	21,3%	19,1%	22,4%	25,5%	19,7%	21,4%	22,9%	22,9%
cuiaba	17,0%	18,8%	19,7%	18,8%	22,1%	20,7%	19,5%	22,2%	22,7%	23,8%	23,2%
curitiba	13,6%	17,8%	16,8%	17,5%	17,7%	18,8%	18,1%	18,7%	18,7%	18,9%	21,3%
florianopolis	13,4%	15,4%	15,6%	16,6%	16,8%	16,2%	16,9%	16,1%	17,3%	18,6%	19,5%
fortaleza	14,8%	17,2%	18,1%	18,3%	17,6%	17,7%	19,5%	19,8%	20,4%	20,9%	20,8%
goiania	13,1%	14,2%	15,1%	15,8%	17,0%	15,6%	13,9%	17,1%	19,6%	18,6%	20,5%
joao pessoa	14,9%	15,2%	16,1%	18,6%	18,6%	17,8%	19,2%	20,9%	19,3%	19,0%	19,2%
macapa	15,6%	17,6%	19,9%	17,5%	20,1%	20,1%	20,3%	20,9%	20,1%	20,0%	21,4%
maceio	12,7%	14,6%	17,1%	17,7%	18,7%	18,9%	21,2%	18,8%	19,9%	17,7%	20,6%
manaus	16,6%	19,2%	20,4%	20,1%	21,0%	20,2%	23,8%	22,7%	23,6%	23,5%	23,7%
natal	14,1%	15,8%	16,2%	19,4%	16,6%	19,3%	18,0%	18,6%	19,1%	20,3%	20,8%
palmas	10,6%	13,2%	14,0%	15,2%	14,8%	16,0%	14,7%	16,1%	16,2%	16,5%	17,2%
porto alegre	14,9%	15,3%	18,4%	18,3%	17,5%	18,7%	19,7%	19,9%	17,4%	20,4%	20,0%
porto velho	17,4%	16,1%	19,0%	19,1%	19,1%	20,5%	20,4%	20,1%	21,0%	22,8%	21,6%
recife	14,1%	16,8%	17,5%	18,2%	19,3%	19,3%	20,4%	19,6%	19,4%	19,7%	21,4%
rio branco	16,0%	18,4%	18,6%	21,1%	19,5%	21,3%	22,6%	23,4%	21,9%	21,9%	21,9%
rio de janeiro	16,0%	16,9%	18,3%	19,9%	19,6%	19,3%	17,0%	19,6%	22,1%	21,6%	21,1%
salvador	14,0%	13,5%	14,9%	15,5%	15,0%	18,2%	17,3%	19,7%	19,7%	19,1%	19,0%
sao luis	11,4%	11,8%	13,1%	13,2%	14,2%	15,0%	15,6%	15,6%	16,5%	15,4%	15,4%
sao paulo	14,7%	15,8%	16,5%	18,6%	19,6%	17,7%	20,4%	19,6%	18,3%	21,5%	21,0%
teresina	12,1%	13,5%	12,7%	15,0%	14,9%	15,3%	16,3%	15,7%	16,0%	17,2%	18,2%
vitoria	12,5%	15,2%	15,0%	16,2%	15,9%	17,1%	15,8%	16,5%	17,1%	18,3%	18,0%
distrito federal	11,1%	12,7%	15,6%	15,5%	16,3%	17,8%	15,2%	16,5%	15,8%	17,7%	17,4%


```

SELECT ANO
      , CIDADE
      , SUM(VÁLIDOS) AS VÁLIDOS
      , SUM(SOBREPESO) AS SOBREPESO
      , SUM(OBESIDADE) AS OBESIDADE
FROM (SELECT ANO
      , CIDADE
      , (CASE WHEN (IMC > 0 AND IMC < 100) THEN 1 ELSE 0 END) AS VÁLIDOS
      , (CASE WHEN (IMC >= 25 AND IMC < 30) THEN 1 ELSE 0 END) AS SOBREPESO
      , (CASE WHEN (IMC >= 30 AND IMC < 100) THEN 1 ELSE 0 END) AS OBESIDADE
FROM VIGITEL) V
GROUP BY ANO, CIDADE
ORDER BY ANO, CIDADE

```

Ranking dos locais, 2009 x 2019

Pos	Cidade	2009
1	porto velho	17,4%
2	cuiaba	17,0%
3	campo grande	16,6%
4	manaus	16,6%
5	rio branco	16,0%
6	rio de janeiro	16,0%
7	macapa	15,6%
8	aracaju	15,5%
9	joao pessoa	14,9%
10	porto alegre	14,9%
11	fortaleza	14,8%
12	sao paulo	14,7%
13	boa vista	14,6%
14	recife	14,1%
15	natal	14,1%
16	salvador	14,0%
17	curitiba	13,6%
18	florianopolis	13,4%
19	goiania	13,1%
20	belem	13,1%
21	maceio	12,7%
22	vitoria	12,5%
23	belo horizonte	12,5%
24	teresina	12,1%
25	sao luis	11,4%
26	distrito federal	11,1%
27	palmas	10,6%

Pos	Cidade	2019
1	manaus	23,7%
2	cuiaba	23,2%
3	campo grande	22,9%
4	boa vista	22,7%
5	rio branco	21,9%
6	porto velho	21,6%
7	recife	21,4%
8	macapa	21,4%
9	curitiba	21,3%
10	rio de janeiro	21,1%
11	sao paulo	21,0%
12	natal	20,8%
13	fortaleza	20,8%
14	aracaju	20,8%
15	maceio	20,6%
16	goiania	20,5%
17	belem	20,1%
18	porto alegre	20,0%
19	belo horizonte	19,8%
20	florianopolis	19,5%
21	joao pessoa	19,2%
22	salvador	19,0%
23	teresina	18,2%
24	vitoria	18,0%
25	distrito federal	17,4%
26	palmas	17,2%
27	sao luis	15,4%

Pos	Cidade	evol
1	boa vista	8,1%
2	maceio	7,9%
3	curitiba	7,7%
4	goiania	7,4%
5	belo horizonte	7,3%
6	recife	7,3%
7	manaus	7,2%
8	belem	7,0%
9	natal	6,8%
10	palmas	6,6%
11	campo grande	6,3%
12	cuiaba	6,3%
13	distrito federal	6,3%
14	sao paulo	6,3%
15	florianopolis	6,1%
16	teresina	6,0%
17	fortaleza	5,9%
18	rio branco	5,9%
19	macapa	5,8%
20	vitoria	5,5%
21	aracaju	5,3%
22	porto alegre	5,2%
23	rio de janeiro	5,2%
24	salvador	5,0%
25	joao pessoa	4,2%
26	porto velho	4,2%
27	sao luis	4,0%

Todas as cidades, e a capital federal, tiveram aumento no índice de **obesidade** entre 2009 e 2019. Em 2019, o maior índice de **obesidade** foi da cidade de manaus (23,7%) e o menor da cidade de são luis (15,4%). O maior aumento em pontos percentuais (valor **evol**) ocorreu na cidade de boa vista (8,1%) e o menor na cidade de são luis (4,0%).

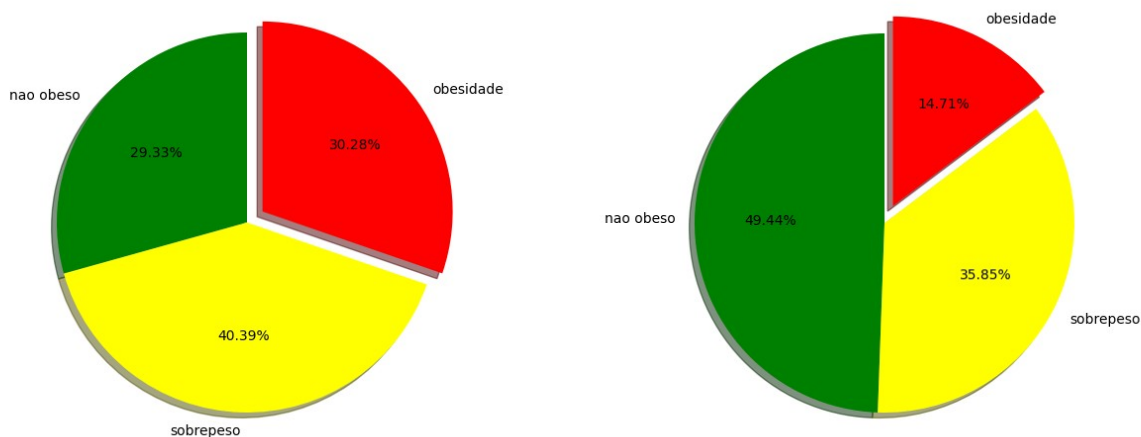
4.4 Relação pressão alta e obesidade (2019)

Será que a **obesidade** impacta na **pressão alta**?

```
SELECT SUM(NAO_OBESO) AS NAO_OBESO
      , SUM(SOBREPESO) AS SOBPES0
      , SUM(OBESIDADE) AS OBESIDADE
FROM(
SELECT (CASE WHEN (IMC > 0 AND IMC < 25) THEN 1 ELSE 0 END) AS NAO_OBESO
      , (CASE WHEN (IMC >= 25 AND IMC < 30) THEN 1 ELSE 0 END) AS SOBPES0
      , (CASE WHEN (IMC >= 30 AND IMC < 100) THEN 1 ELSE 0 END) AS OBESIDADE
FROM VIGITEL
WHERE IMC > 0 AND IMC < 100 AND ANO = 2019 AND PRESSAO_ALTA = 1) V
```

```
SELECT SUM(NAO_OBESO) AS NAO_OBESO
      , SUM(SOBREPESO) AS SOBPES0
      , SUM(OBESIDADE) AS OBESIDADE
FROM(
SELECT (CASE WHEN (IMC > 0 AND IMC < 25) THEN 1 ELSE 0 END) AS NAO_OBESO
      , (CASE WHEN (IMC >= 25 AND IMC < 30) THEN 1 ELSE 0 END) AS SOBPES0
      , (CASE WHEN (IMC >= 30 AND IMC < 100) THEN 1 ELSE 0 END) AS OBESIDADE
FROM VIGITEL
WHERE IMC > 0 AND IMC < 100 AND ANO = 2019 AND PRESSAO_ALTA = 2) V
```

COM pressão alta x SEM pressão alta



A resposta a pergunta sobre o impacto da **obesidade** depende de pesquisas de outras áreas, como a biomédica, mas os números mostram uma forte relação entre a **obesidade/sobrepeso** e **pressão alta**. Nos indivíduos que disseram não possuir pressão alta, quase a metade possui IMC **normal** (49,44%). Já nos indivíduos que declararam pressão alta, a **obesidade** é o dobro dos que não possuem pressão alta, 30,28% contra 14,71%.

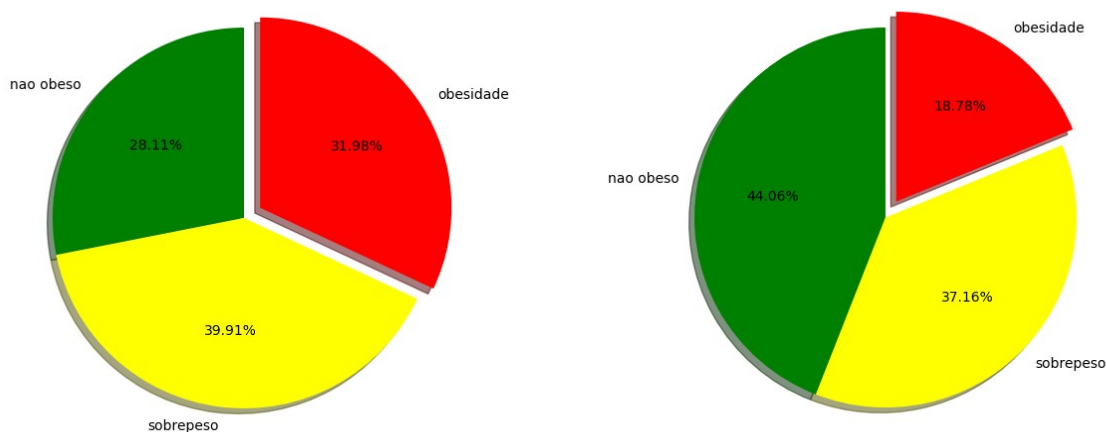
4.5 Relação diabetes e obesidade (2019)

Será que a **obesidade** impacta na **diabetes**?

```
SELECT SUM(NAO_OBESO) AS NAO_OBESO
      , SUM(SOBREPESO) AS SOBPESO
      , SUM(OBESIDADE) AS OBESIDADE
FROM(
SELECT (CASE WHEN (IMC > 0 AND IMC < 25) THEN 1 ELSE 0 END) AS NAO_OBESO
      , (CASE WHEN (IMC >= 25 AND IMC < 30) THEN 1 ELSE 0 END) AS SOBPESO
      , (CASE WHEN (IMC >= 30 AND IMC < 100) THEN 1 ELSE 0 END) AS OBESIDADE
FROM VIGITEL
WHERE IMC > 0 AND IMC < 100 AND ANO = 2019 AND DIABETES = 1) V
```

```
SELECT SUM(NAO_OBESO) AS NAO_OBESO
      , SUM(SOBREPESO) AS SOBPESO
      , SUM(OBESIDADE) AS OBESIDADE
FROM(
SELECT (CASE WHEN (IMC > 0 AND IMC < 25) THEN 1 ELSE 0 END) AS NAO_OBESO
      , (CASE WHEN (IMC >= 25 AND IMC < 30) THEN 1 ELSE 0 END) AS SOBPESO
      , (CASE WHEN (IMC >= 30 AND IMC < 100) THEN 1 ELSE 0 END) AS OBESIDADE
FROM VIGITEL
WHERE IMC > 0 AND IMC < 100 AND ANO = 2019 AND DIABETES = 2) V
```

COM diabetes x SEM diabetes



Assim como na **pressão alta**, a resposta a pergunta sobre o impacto da **obesidade** no **diabetes** depende de pesquisas de outras áreas, aqui os números também mostram uma forte relação entre a **obesidade/sobrepeso** e o **diabetes**. Nos indivíduos que disseram não possuir diabetes, os que possuem IMC **normal** é também quase a metade (44,06%). Já nos indivíduos que declararam diabetes, a **obesidade** é quase o dobro dos que não possuem diabetes, 31,98% contra 18,78%.

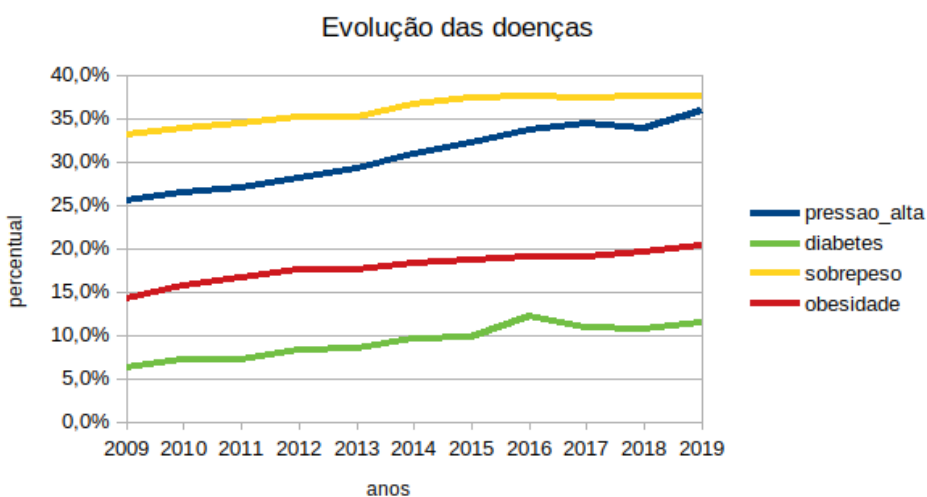
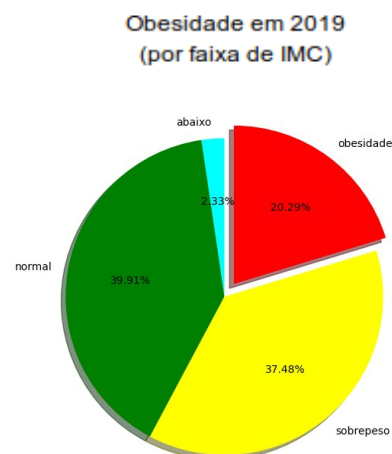
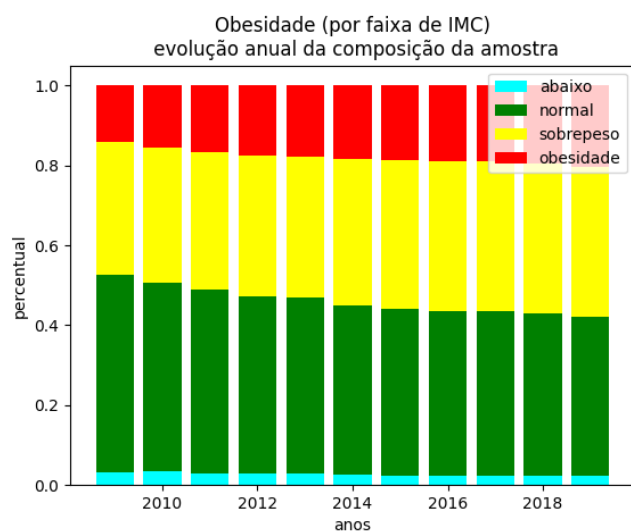
Olhando somente os números, a impressão é que as duas doenças analisadas aqui, **pressão alta** e **diabetes**, possuem forte relação com a **obesidade**, mas a maior relação entre as duas é a **pressão arterial**. A confirmação depende de outras áreas de conhecimento.

5. Apresentação dos Resultados

5.1 Workflow

1 – Problem Statement	2 – Outcomes/Predictions	3 – Data Acquisition
A análise da obesidade no Brasil, sua evolução anual e relação com outras doenças, como pressão alta e diabetes, impacta nas decisões e políticas públicas de saúde.	Espera-se confirmar o aumento da obesidade e a relação com outras doenças.	Dados da Pesquisa de Vigilância de Fatores de Risco e Proteção para Doenças crônicas por Inquérito Telefônico (Vigitel) do Ministério da Saúde.
4 – Modeling	5 – Model Evaluation	6 – Data Preparation
Análise exploratória dos dados.	Confronto com estudos de outras áreas, como a biomedicina.	Selecionando dados da coleta do vigitel, como valor de IMC, ano e cidade com a tabela de IMC (Índice de Massa Corporal).

5.2 Evolução das doenças



6. Links

Estes são os *links* para os arquivos deste TCC.

6.1 Vídeo de apresentação

Análise Vigitel (obesidade)

<https://youtu.be/6PHaSqzYzow>

6.2 Repositório do Projeto

<https://github.com/mleal-dev/PUCMinas.git>

Conteúdo:

TCC_Analise_Vigitel.pdf	TCC (este documento)
Analise_Vigitel.ipynb	Arquivo <i>jupyter notebook</i> com as consultas, tabelas e gráficos
vigitel.py	Arquivo <i>python</i> para rotinas de importação e consulta de dados
DB_VIGITEL.db	Arquivo <i>SQLite</i> com dados importados
calculo_aux.xlsx	Planilha auxiliar para a apresentação de tabelas e gráficos

REFERÊNCIAS

(1) Detalhes do **Vigitel** em:

<https://saude.gov.br/saude-de-a-z/vigitel>

<https://www.saude.gov.br/noticias/agencia-saude/46191-ministerio-da-saude-quer-conhecer-o-perfil-da-saude-dos-brasileiros>

(2) O último formulário usado na pesquisa **Vigitel** pode ser encontrado neste documento:

<https://portalarquivos.saude.gov.br/images/pdf/2020/Abril/27/vigitel-brasil-2019-vigilancia-fatores-risco.pdf>

(3) IMC:

https://pt.wikipedia.org/wiki/Índice_de_massa_corporal

(4) **kg** (quilograma) é uma unidade de massa:

<https://pt.wikipedia.org/wiki/Quilograma>

(5) Os dados do **Vigitel** podem ser obtidos em:

<http://svs.aids.gov.br/download/Vigitel/>

http://svs.aids.gov.br/bases_vigitel_viva/vigitel.php

(6) A descrição de todas as variáveis disponíveis para tabulação pode ser encontrada em

<http://tabnet.datasus.gov.br/cgi/vigitel/vigteldescr.htm>