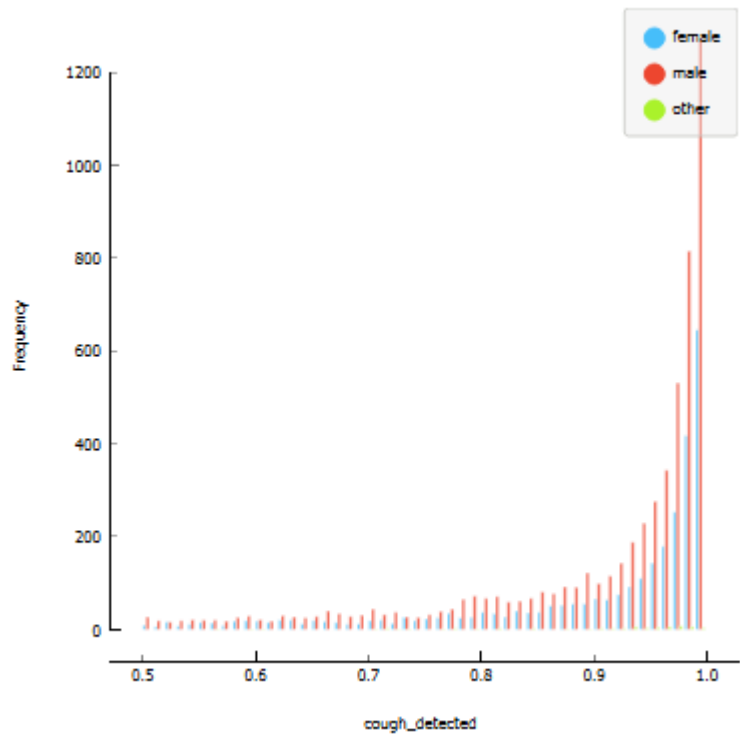


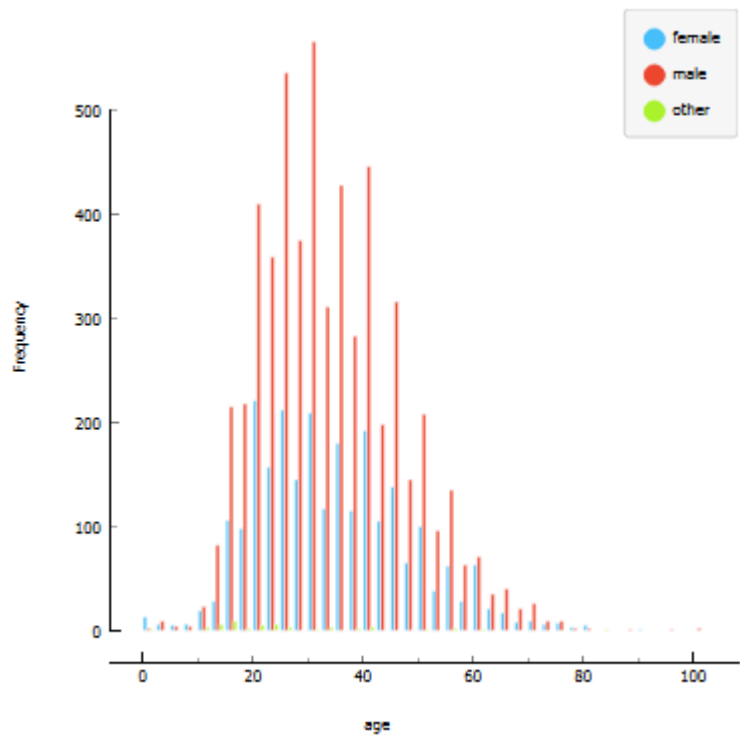
Distribution of 'cough_detected' with columns split by 'status'

Para todas as análises e gráficos a seguir, selecionamos apenas as amostras de áudio com detecção de tosse maior que 50% ($cough_detected > 0,5$). O gráfico acima mostra a distribuição das amostras divididas de acordo com a classificação do estado de saúde indicada pelo próprio indivíduo (variável "status"): "COVID-19" = com COVID-19; "healthy" = saudável; "symptomatic" = com sintomas para alguma condição clínica.



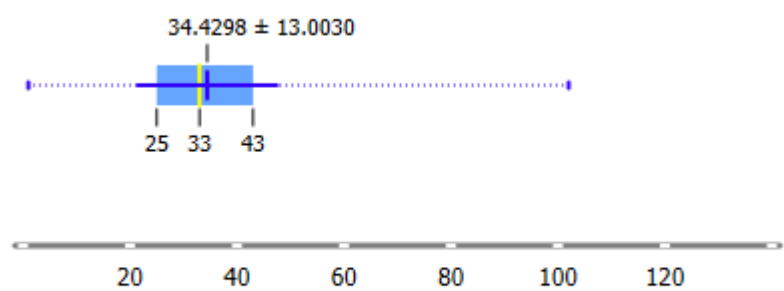
Distribution of 'cough_detected' with columns split by 'gender'

O gráfico acima mostra a distribuição das amostras de áudio separadas de acordo com o gênero dos indivíduos. É possível observar uma maior proporção de indivíduos do gênero masculino (barra vermelha) do que feminino (barra azul).



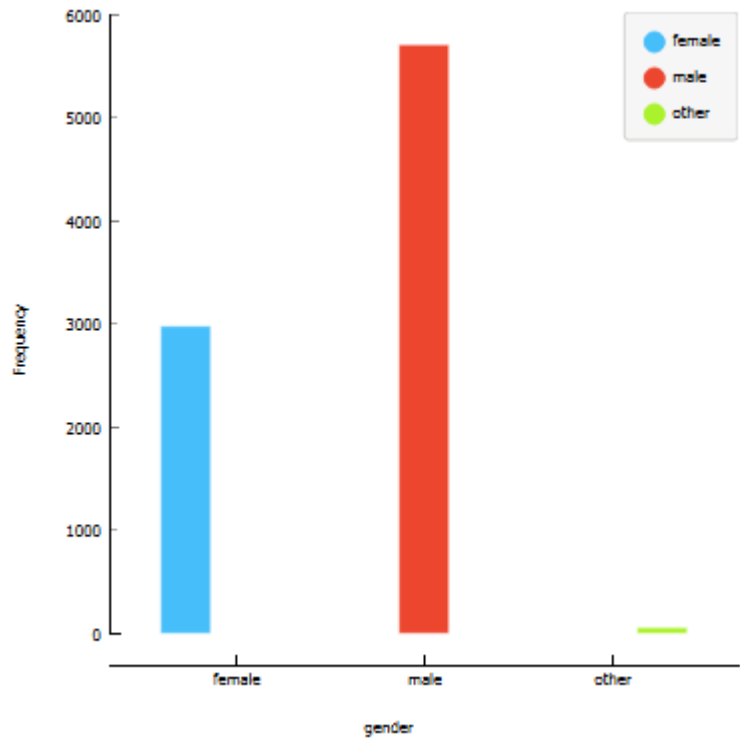
Distribution of 'age' with columns split by 'gender'

O gráfico acima mostra a distribuição de idade dos indivíduos avaliados, separados de acordo o gênero. Também é possível observar maior proporção de avaliados do gênero masculino (barra vermelha) do que feminino (barra azul).



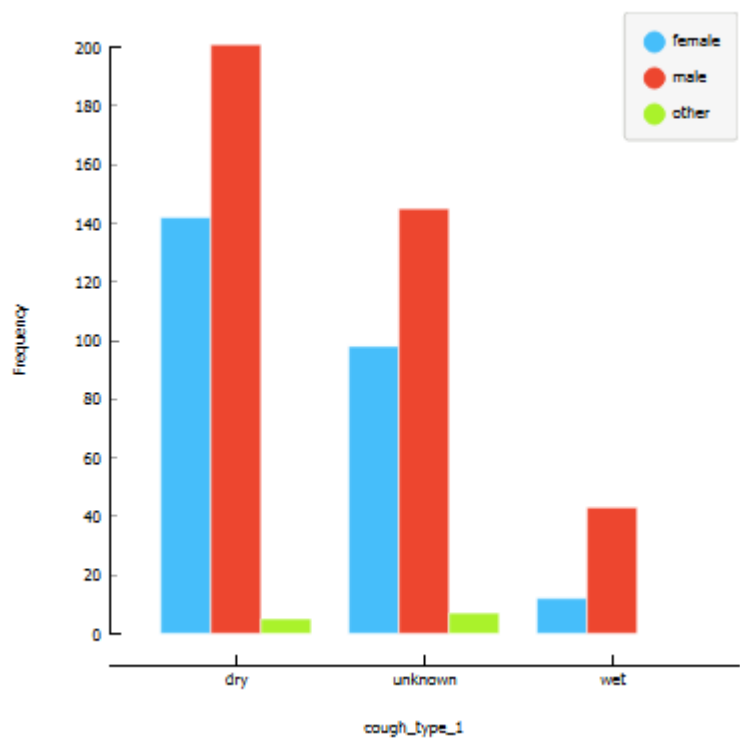
Box plot for attribute 'age'

O Box Plot acima mostra o intervalo de idade dos indivíduos (variável "age"). A mediana corresponde a 33 anos, a média é 34.4 anos, o primeiro quartil é 25 anos e o terceiro quartil é 43 anos.



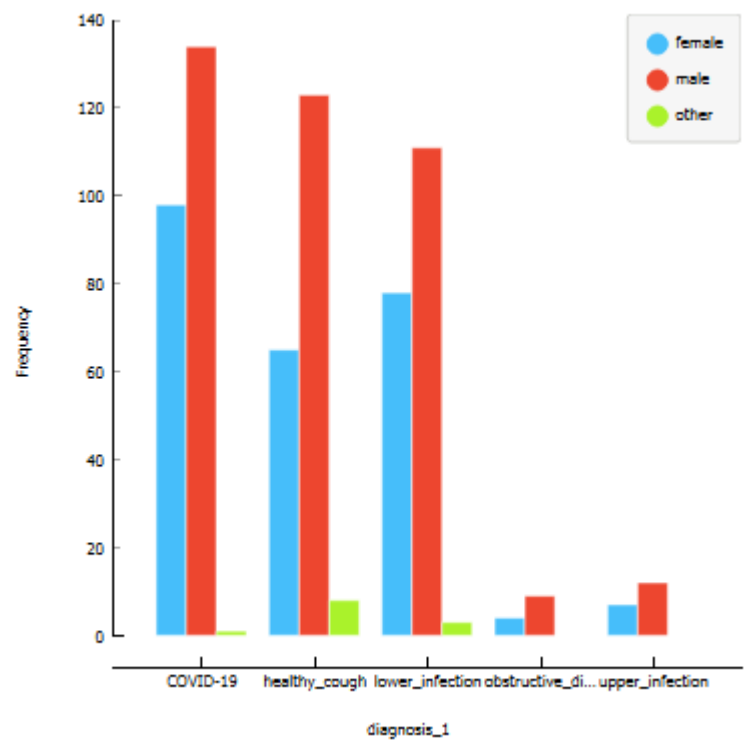
Distribution of 'gender' with columns split by 'gender'

Distribuição dos indivíduos avaliados de acordo com o gênero - masculino ou feminino. Há maior proporção de indivíduos do gênero masculino do que feminino. Alguns indivíduos optaram por não definir o gênero ("other" = outro). Essa análise da base de dados da seleção de amostras com detecção de tosse maior que 50% está semelhante à análise da base de dados original com todas as amostras.



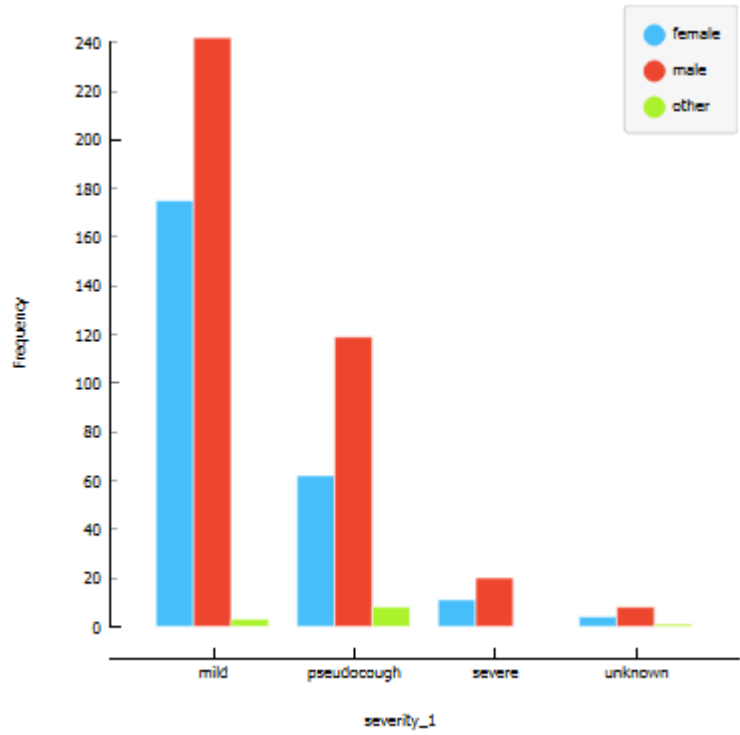
Distribution of 'cough_type_1' with columns split by 'gender'

Segundo a avaliação do especialista 1, o gráfico mostra como as amostras de áudio foram classificadas conforme os três tipos de tosse: "dry" = seca; "unknown" = não definida; "wet" = úmida).



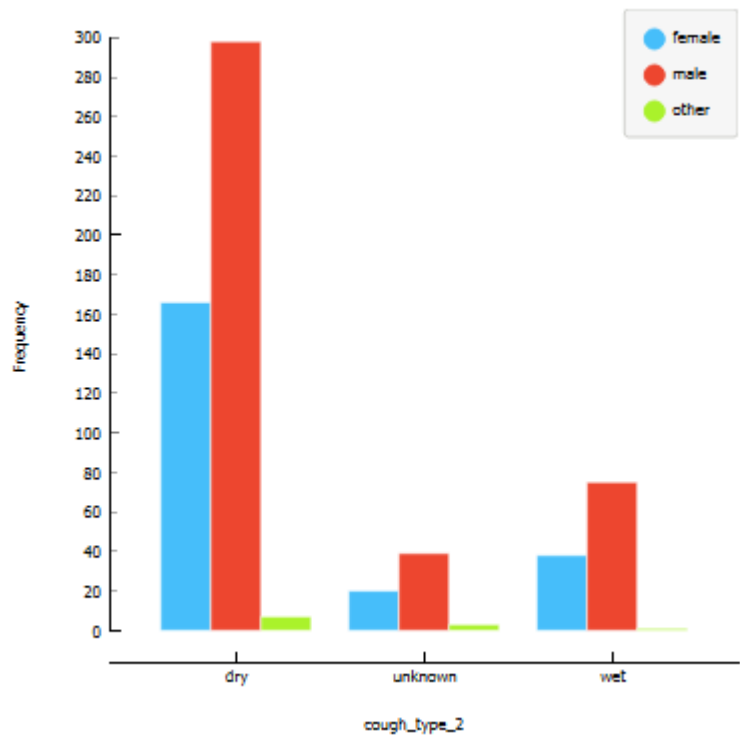
Distribution of 'diagnosis_1' with columns split by 'gender'

Segundo a avaliação do especialista 1, o gráfico mostra a distribuição dos tipos de diagnóstico que caracterizam o áudio da tosse de cada indivíduo. Os possíveis diagnósticos foram: "COVID-19" = com Covid-19, "healthy_cough" = com tosse saudável, "lower_infection" = com infecção do trato respiratório inferior, "obstructive_disease" = com doença respiratória obstrutiva, "upper_infection" = com infecção do trato respiratório superior.



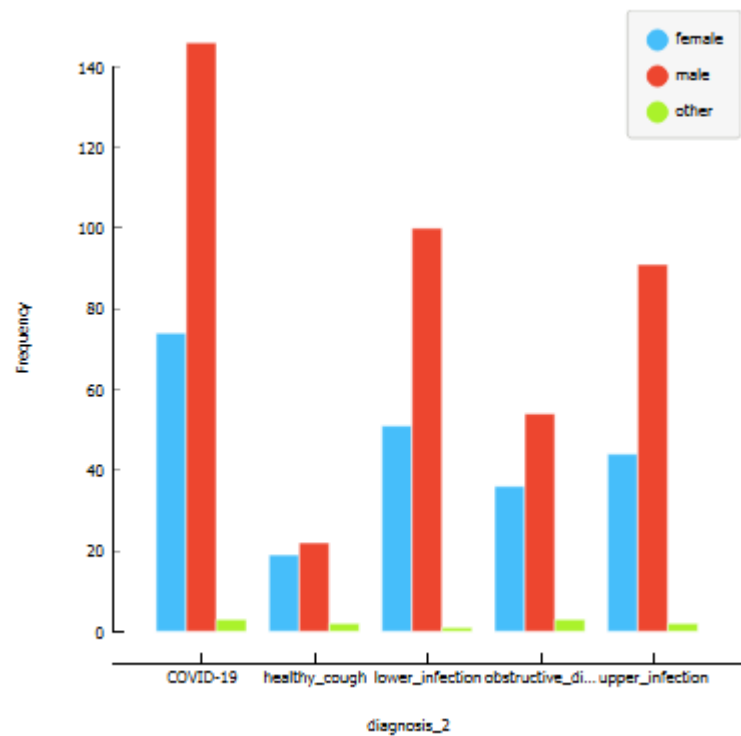
Distribution of 'severity_1' with columns split by 'gender'

Segundo a avaliação do especialista 1, o gráfico mostra a distribuição dos tipos de severidade que caracterizam a condição clínica do indivíduo. Os quadros de severidade analisados foram: "mild" = leve, "pseudocough" = pseudo-tosse, "severe" = severa, "unknown" = não determinada.



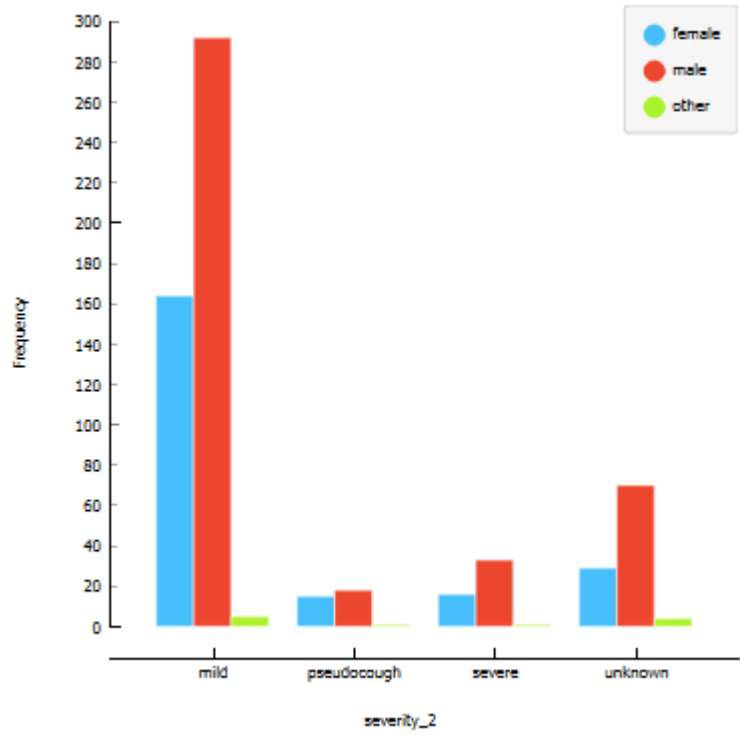
Distribution of 'cough_type_2' with columns split by 'gender'

Segundo a avaliação do especialista 2, o gráfico mostra como as amostras de áudio foram classificadas conforme os três tipos de tosse: "dry" = seca; "unknown" = não definida; "wet" = úmida).



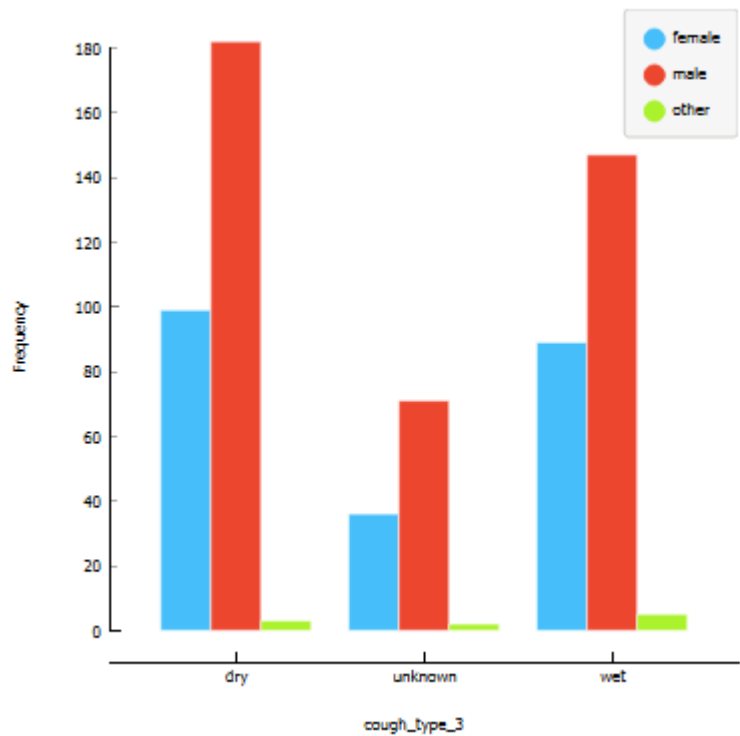
Distribution of 'diagnosis_2' with columns split by 'gender'

Segundo a avaliação do especialista 2, o gráfico mostra a distribuição dos tipos de diagnóstico que caracterizam o áudio da tosse de cada indivíduo. Os possíveis diagnósticos foram: "COVID-19" = com Covid-19, "healthy_cough" = com tosse saudável, "lower_infection" = com infecção do trato respiratório inferior, "obstructive_disease" = com doença respiratória obstrutiva, "upper_infection" = com infecção do trato respiratório superior.



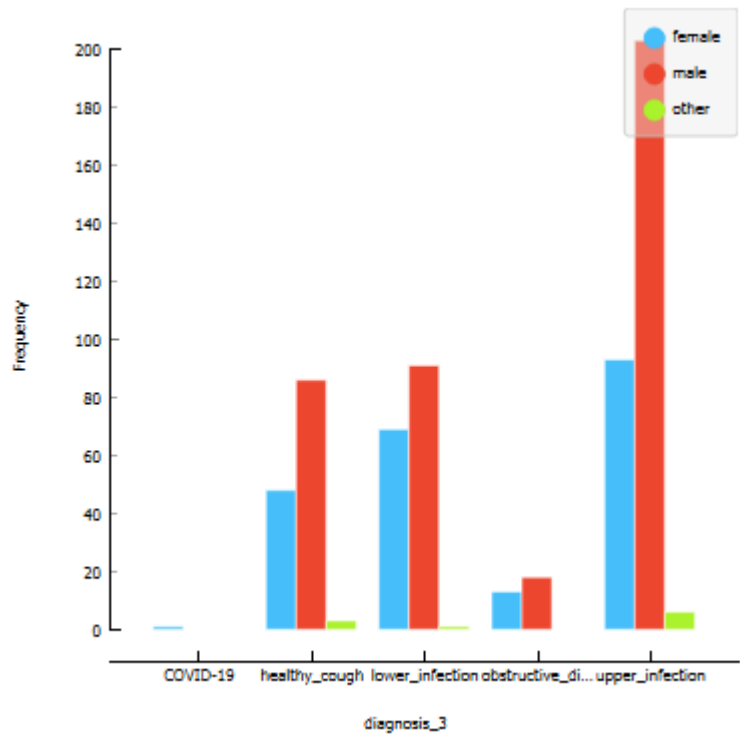
Distribution of 'severity_2' with columns split by 'gender'

Segundo a avaliação do especialista 2, o gráfico mostra a distribuição dos tipos de severidade que caracterizam a condição clínica do indivíduo. Os quadros de severidade analisados foram: "mild" = leve, "pseudocough" = pseudo-tosse, "severe" = severa, "unknown" = não determinada.



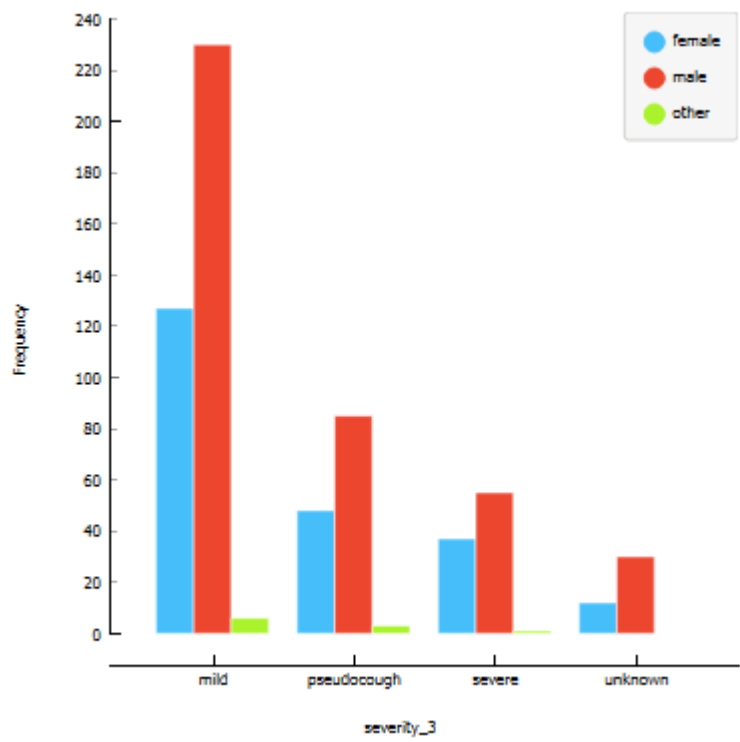
Distribution of 'cough_type_3' with columns split by 'gender'

Segundo a avaliação do especialista 3, o gráfico mostra como as amostras de áudio foram classificadas conforme os três tipos de tosse: "dry" = seca; "unknown" = não definida; "wet" = úmida).



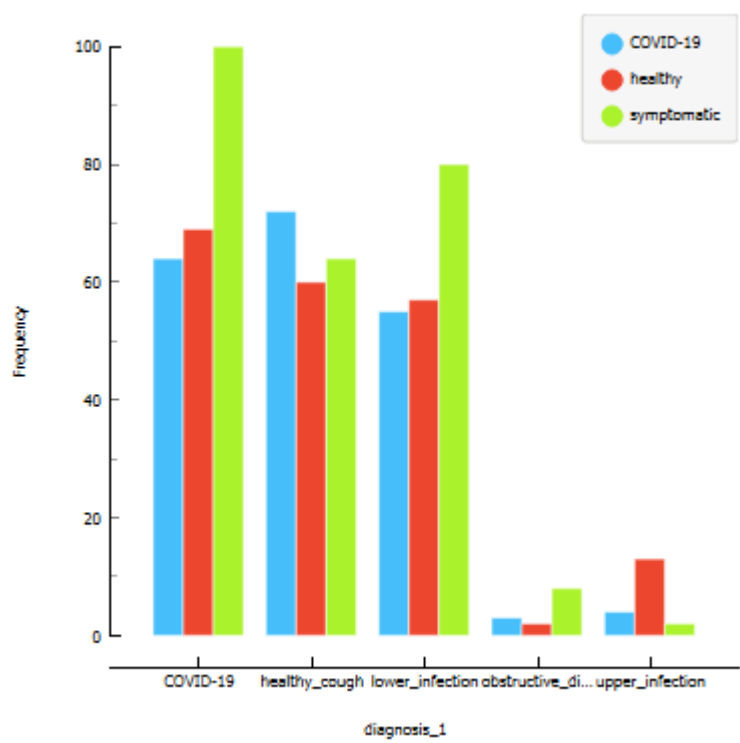
Distribution of 'diagnosis_3' with columns split by 'gender'

Segundo a avaliação do especialista 3, o gráfico mostra a distribuição dos tipos de diagnóstico que caracterizam o áudio da tosse de cada indivíduo. Os possíveis diagnósticos foram: "COVID-19" = com Covid-19, "healthy_cough" = com tosse saudável, "lower_infection" = com infecção do trato respiratório inferior, "obstructive_disease" = com doença respiratória obstrutiva, "upper_infection" = com infecção do trato respiratório superior.



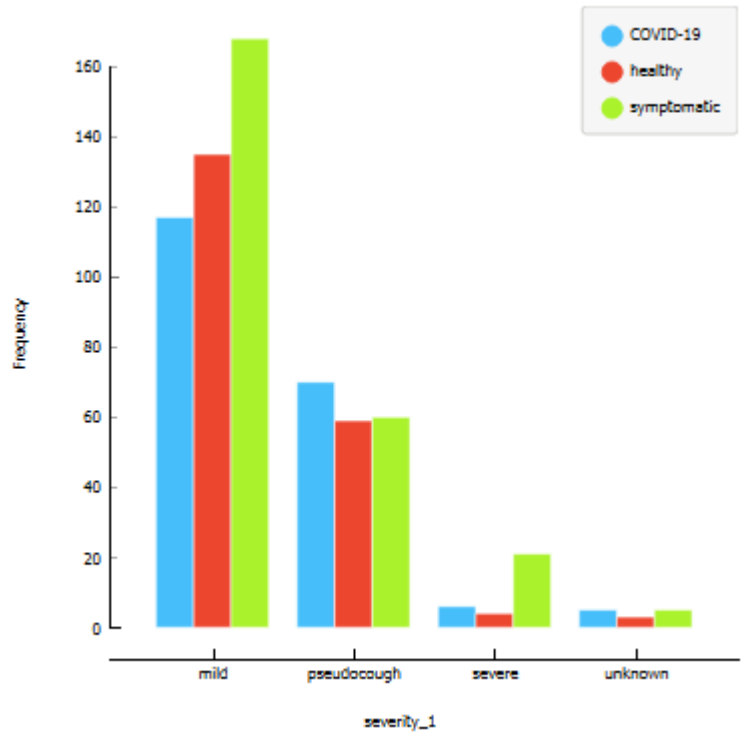
Distribution of 'severity_3' with columns split by 'gender'

Segundo a avaliação do especialista 3, o gráfico mostra a distribuição dos tipos de severidade que caracterizam a condição clínica do indivíduo. Os quadros de severidade analisados foram: "mild" = leve, "pseudocough" = pseudo-tosse, "severe" = severa, "unknown" = não determinada.



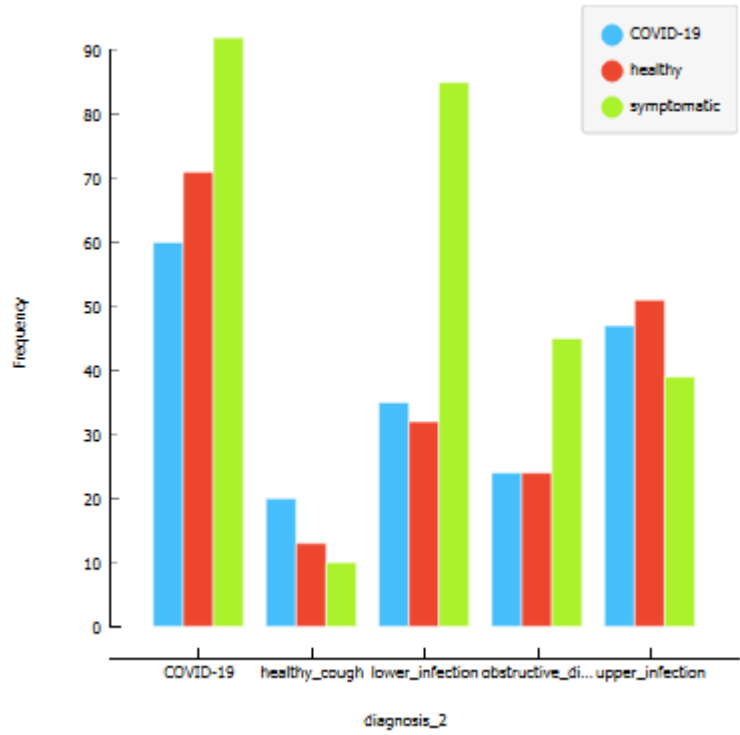
Distribution of 'diagnosis_1' with columns split by 'status'

O gráfico compara as classes de diagnósticos conforme avaliação do especialista 1 com a própria avaliação do indivíduo (variável "status"). É possível notar diferenças entre o diagnóstico da tosse dado pelo especialista 1 e o "status" definido pelo próprio indivíduo. Por exemplo, para o "status" de condição de Covid-19 (barra azul), alguns áudios de tosse foram diagnosticados como tosse saudável ("healthy_cough") ou como outras infecções do trato respiratório.



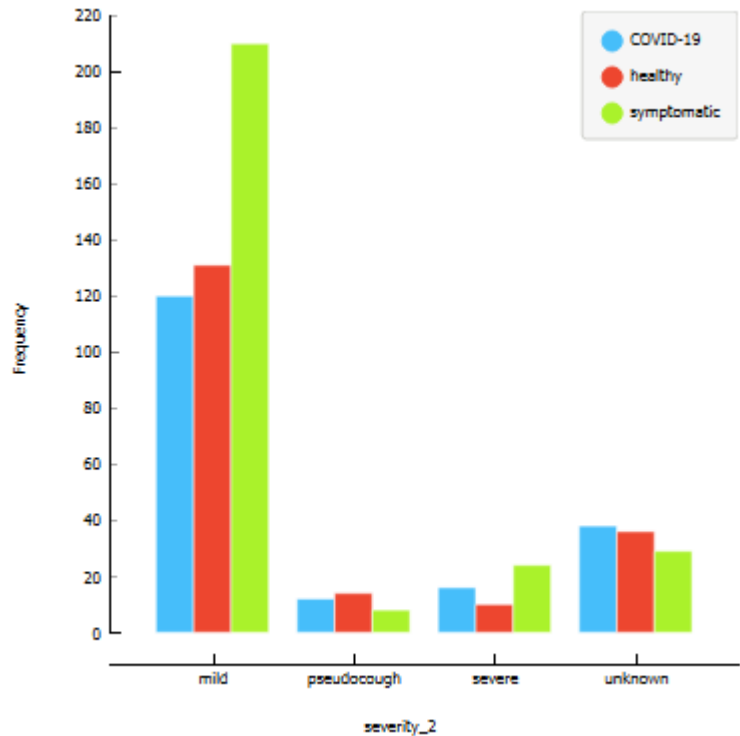
Distribution of 'severity_1' with columns split by 'status'

O gráfico mostra a distribuição das amostras em relação à severidade da condição clínica, avaliada pelo especialista 1, separado conforme o estado de saúde definido pelo indivíduo ("status"). É possível observar que a maior parte das avaliações foram leves (mild).



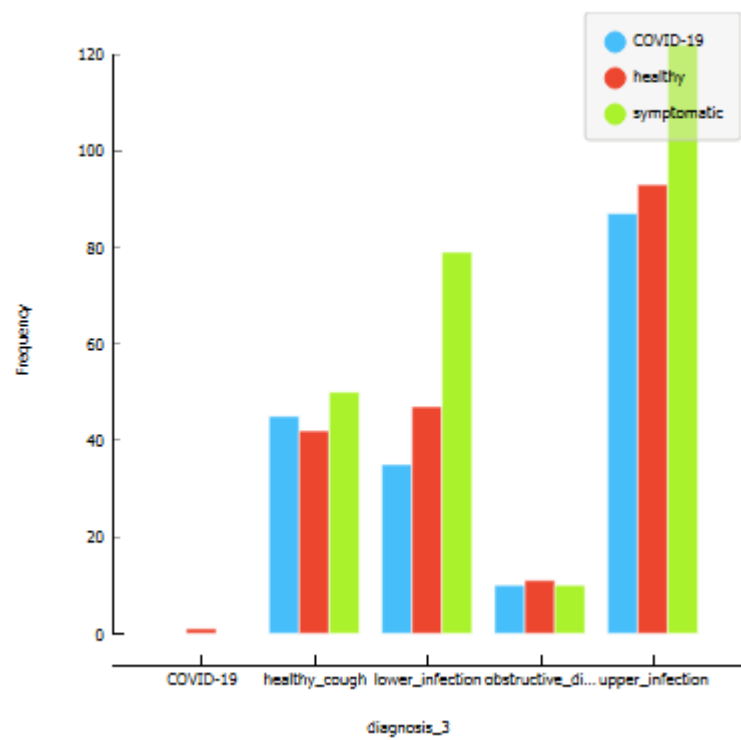
Distribution of 'diagnosis_2' with columns split by 'status'

O gráfico compara as classes de diagnósticos conforme avaliação do especialista 2 com a própria avaliação do indivíduo (variável "status"). É possível notar diferenças entre o diagnóstico da tosse dado pelo especialista 2 e o "status" definido pelo próprio indivíduo. Por exemplo, para o "status" de condição de Covid-19 (barra azul), alguns áudios de tosse foram diagnosticados como tosse saudável ("healthy_cough") ou como outras infecções do trato respiratório.



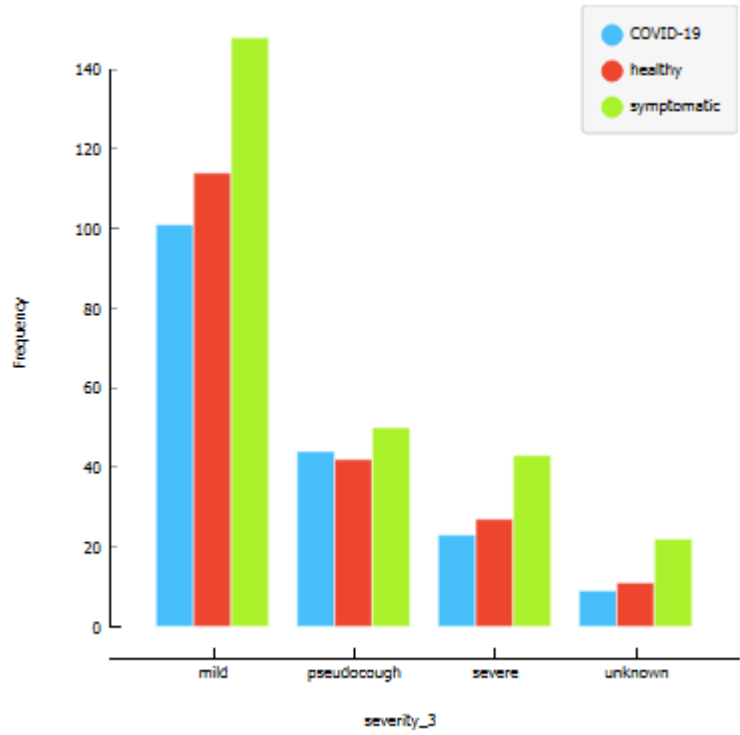
Distribution of 'severity_2' with columns split by 'status'

O gráfico mostra a distribuição das amostras em relação à severidade da condição clínica, avaliada pelo especialista 2, separado coforme a classificação do estado de saúde definido pelo indivíduo ("status"). É possível observar que a maior parte das avaliações foram leves (mild).



Distribution of 'diagnosis_3' with columns split by 'status'

O gráfico compara as classes de diagnósticos conforme avaliação do especialista 3 com a própria avaliação do indivíduo (variável "status"). É possível notar diferenças entre o diagnóstico da tosse dado pelo especialista 3 e o "status" definido pelo próprio indivíduo. Por exemplo, para o "status" de condição de Covid-19 (barra azul), alguns áudios de tosse foram diagnosticados como tosse saudável ("healthy_cough") ou como outras infecções do trato respiratório.



Distribution of 'severity_3' with columns split by 'status'

O gráfico mostra a distribuição das amostras em relação à severidade da condição clínica, avaliada pelo especialista 3, separado conforme a classificação do estado de saúde definido pelo indivíduo ("status"). É possível observar que a maior parte das avaliações foram leves (mild).

Data instances: 11
Features: 4
Meta attributes: 1

	name	mild	pseudocough	severe	unknown
1	intercept	1.94906	-0.767292	0.193123	-1.37489
2	dyspnea_1=False	0.695512	0.0657825	-0.859512	0.0982171
3	dyspnea_1=True	-0.69538	-0.0656971	0.85961	-0.0985325
4	wheezing_1=False	-0.12712	0.63879	-0.167478	-0.344191
5	wheezing_1=True	0.127252	-0.638704	0.167577	0.343876
6	choking_1=False	-0.129536	0.0346195	0.0751384	0.019778
7	choking_1=True	0.129668	-0.0345341	-0.0750403	-0.0200933
8	congestion_1=False	-0.478673	0.136934	0.255491	0.0862485
9	congestion_1=True	0.478805	-0.136848	-0.255393	-0.0865639
10	nothing_1=False	0.190717	-0.909461	0.56217	0.156575
11	nothing_1=True	-0.190585	0.909546	-0.562071	-0.15689

A regressão logística foi realizada para verificar qual característica do áudio da tosse ("dyspnea", "wheezing", "choking", "congestion" e "nothing") mais influenciou para o especialista determinar uma condição clínica grave ("severe"). Para o especialista 1, o fato do indivíduo apresentar "dyspnea" = dispneia (dificuldade para respirar) foi o que mais interferiu na avaliação, apresentando coeficiente de 0.85.

Data instances: 13
Features: 4
Meta attributes: 1

	name	mild	pseudocough	severe	unknown
1	intercept	2.40692	-1.73705	1.52491	-2.19479
2	dyspnea_2=False	-0.13983	0.514377	-0.928098	0.553552
3	dyspnea_2=True	0.139111	-0.515014	0.927389	-0.551487
4	wheezing_2=False	-0.609826	0.367279	-0.406295	0.648842
5	wheezing_2=True	0.609107	-0.367915	0.405586	-0.646777
6	stridor_2=False	-0.12148	0.31041	-0.393402	0.204473
7	stridor_2=True	0.120761	-0.311047	0.392693	-0.202408
8	choking_2=False	0.269015	0.232979	-0.981247	0.479253
9	choking_2=True	-0.269734	-0.233616	0.980537	-0.477188
10	congestion_2=False	-0.0621967	-0.525394	0.0192817	0.568309
11	congestion_2=True	0.0614776	0.524758	-0.0199907	-0.566245
12	nothing_2=False	0.0191967	-0.300818	0.444184	-0.162562
13	nothing_2=True	-0.0199158	0.300182	-0.444893	0.164627

A regressão logística foi realizada para verificar qual característica do áudio da tosse ("dyspnea", "wheezing", "choking", "congestion" e "nothing") mais influenciou para o especialista determinar uma condição clínica grave ("severe"). Para o especialista 2, o fato do indivíduo apresentar "choking" = sensação de asfixia, foi o que mais interferiu na avaliação, apresentando coeficiente de 0.98.

Data instances: 11
Features: 4
Meta attributes: 1

	name	mild	pseudocough	severe	unknown
1	intercept	0.896785	-0.799515	0.528307	-0.625577
2	dyspnea_3=False	0.421362	0.307417	-0.310434	-0.418346
3	dyspnea_3=True	-0.423753	-0.306454	0.313372	0.416835
4	wheezing_3=False	0.438331	0.261946	-0.801578	0.101302
5	wheezing_3=True	-0.440721	-0.260982	0.804516	-0.102813
6	stridor_3=False	-0.440843	0.174691	0.144436	0.121716
7	stridor_3=True	0.438453	-0.173727	-0.141498	-0.123227
8	congestion_3=False	-0.685367	0.275978	0.22404	0.185349
9	congestion_3=True	0.682977	-0.275015	-0.221102	-0.18686
10	nothing_3=False	-0.495773	-0.027829	0.14645	0.377153
11	nothing_3=True	0.493383	0.0287925	-0.143512	-0.378664

A regressão logística foi realizada para verificar qual característica do áudio da tosse ("dyspnea", "wheezing", "choking", "congestion" e "nothing") mais influenciou para o especialista determinar uma condição clínica grave ("severe"). Para o especialista 3, o fato do indivíduo apresentar "wheezing" = respiração ofegante, foi o que mais interferiu na avaliação, apresentando coeficiente de 0.80.