



TESTE TÉCNICO – CIENTISTA DE DADOS A3DATA

ANTÔNIO DOS SANTOS RAMOS NETO

1) APRESENTAÇÃO DO DESAFIO

Olá meu nome é Antônio, sou cientista de dados e ao longo desse desafio você irá imergir em um estudo relacionado as ocorrências aeronáuticas na aviação civil brasileira.

Link do repositório: <https://github.com/tonylincon1/Analise-Exploratoria-Dados-CENIPA>

Sobre os Dados

A base de dados de ocorrências aeronáuticas é gerenciada pelo Centro de Investigação e Prevenção de Acidentes Aeronáuticos (CENIPA). Constam nesta base de dados as ocorrências aeronáuticas notificadas ao CENIPA nos últimos 10 anos e que ocorreram em solo brasileiro.

Dentre as informações disponíveis estão os dados sobre as aeronaves envolvidas, fatalidades, local, data, horário dos eventos e informações taxonômicas típicas das investigações de acidentes (AIG). São resguardadas a privacidade de pessoas físicas/jurídicas envolvidas conforme previsto pela Lei de Acesso à Informação (Lei nº 12.527, de 18 de novembro de 2011).

Essa base de dados é composta por informações preliminares provenientes do formulário CENIPA-05 (Ficha de Notificação de Ocorrências Aeronáuticas) e consolidada a partir dos relatórios de investigações publicados. Outra forma de visualização desses dados é pelo Painel SIPAER disponível na página do CENIPA.

Dados proveniente dos Programas de Gerenciamento de Prevenção administrados pelo CENIPA (Emissão de Raio Laser e Risco Baloeiro) não constam nesta base. Esses programas possuem formulários próprios para coleta de dados com foco exclusivo para gerenciamento do risco, enquanto os dados coletados pelo formulário CENIPA-05 têm foco principal na investigação de acidentes (AIG).

Objetivo desse Desafio

O objetivo deste teste é explorar e demonstrar minhas habilidades como Cientista de Dados e criatividade com as análises. Os passos a serem entregues são:

- 1) Apresentação do desafio;
- 2) Explicação do processo utilizado;
- 3) Análise exploratória;
- 4) Hipóteses levantadas;
- 5) Conclusão e insights gerados.

Critérios a Serem Avaliados

- 1) Capacidade analítica;
- 2) Qualidade do storytelling na apresentação do problema e no código;
- 3) Qualidade e clareza do código.

2) EXPLICAÇÃO DO PROCESSO UTILIZADO

2.1) Metodologia Utilizada

A metodologia utilizada nesse estudo se baseia na investigação exploratória das bases de dados afim de descrever os comportamentos e tentar traçar conexões, a linha de segmento do estudo será baseado nas seguintes etapas de investigação:

1) Identificação dos maiores eventos de ocorrência

Investigando por classificação

Investigando por estado

Investigando por cidade

Investigando por hora

2) Investigando o tipo de ocorrência, fator de contribuição e aeronave

Investigando o tipo de ocorrência

Investigando o fator de contribuição

Investigando a aeronave

3) Investigando as recomendações

Essa estratégia de exploração busca entender o fenômeno no item 1), identificar possíveis causas das ocorrências no item 2) e por fim, verificar se as recomendações, item 3), estão sendo efetivas para redução dos acidentes.

2.2) Bibliotecas Utilizadas

```
1 import pandas as pd
2 import seaborn as sns
3 import matplotlib.pyplot as plt
4 import matplotlib.image as mpimg
5 from matplotlib import rcParams
6 from datetime import datetime
7
8 %matplotlib inline
```

2.3) Bases de Dados Utilizadas

As bases de dados utilizadas foram coletadas a partir do portal brasileiro de dados abertos (<https://dados.gov.br/dataset/ocorrencias-aeronauticas-da-aviacao-civil-brasileira>).

2.3.1) Importação das Bases de Dados

```
1 ocorrencias = pd.read_csv("http://sistema.cenipa.aer.mil.br/cenipa/media/opendata/ocorrencia.csv", sep=';')
2 tipo_de_ocorrencia = pd.read_csv("http://sistema.cenipa.aer.mil.br/cenipa/media/opendata/ocorrencia_tipo.csv", sep=';')
3 aeronave = pd.read_csv("http://sistema.cenipa.aer.mil.br/cenipa/media/opendata/aeronave.csv", sep=';')
4 fator_contribuinte = pd.read_csv("http://sistema.cenipa.aer.mil.br/cenipa/media/opendata/fator_contribuinte.csv", sep=';')
5 recomendacao_de_seguranca = pd.read_csv("http://sistema.cenipa.aer.mil.br/cenipa/media/opendata/recomendacao.csv", sep=';')
```

2.3.2) Conhecendo as Bases de Dados

```
1 tipo_de_ocorrencia.info()
2 tipo_de_ocorrencia.head()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 6283 entries, 0 to 6282
Data columns (total 4 columns):
#   Column                                Non-Null Count  Dtype
---  -
0   codigo_ocorrencia1                    6283 non-null   int64
1   ocorrencia_tipo                       6283 non-null   object
2   ocorrencia_tipo_categoria             6283 non-null   object
3   taxonomia_tipo_icao                    6283 non-null   object
dtypes: int64(1), object(3)
memory usage: 196.5+ KB
```

```
1 fator_contribuinte.info()
2 fator_contribuinte.head()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 4485 entries, 0 to 4484
Data columns (total 5 columns):
#   Column                                Non-Null Count  Dtype
---  -
0   codigo_ocorrencia3                    4485 non-null   int64
1   fator_nome                            4485 non-null   object
2   fator_aspecto                         4485 non-null   object
3   fator_condicionante                   4485 non-null   object
4   fator_area                            4485 non-null   object
dtypes: int64(1), object(4)
memory usage: 175.3+ KB
```

```
1 ocorrencias.info()
2 ocorrencias.head()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 6114 entries, 0 to 6113
Data columns (total 22 columns):
#   Column                                Non-Null Count  Dtype
---  -
0   codigo_ocorrencia                     6114 non-null   int64
1   codigo_ocorrencia1                    6114 non-null   int64
2   codigo_ocorrencia2                    6114 non-null   int64
3   codigo_ocorrencia3                    6114 non-null   int64
4   codigo_ocorrencia4                    6114 non-null   int64
5   ocorrencia_classificacao              6114 non-null   object
6   ocorrencia_latitude                   4552 non-null   object
7   ocorrencia_longitude                  4552 non-null   object
8   ocorrencia_cidade                     6114 non-null   object
9   ocorrencia_uf                         6114 non-null   object
10  ocorrencia_pais                       6114 non-null   object
11  ocorrencia_aerodromo                  6114 non-null   object
12  ocorrencia_dia                        6114 non-null   object
13  ocorrencia_hora                       6113 non-null   object
14  investigacao_aeronave_liberada         5827 non-null   object
15  investigacao_status                    5773 non-null   object
16  divulgacao_relatorio_numero            5319 non-null   object
17  divulgacao_relatorio_publicado         6114 non-null   object
18  divulgacao_dia_publicacao              1577 non-null   object
19  total_recomendacoes                   6114 non-null   int64
20  total_aeronaves_envolvidas             6114 non-null   int64
21  ocorrencia_saida_pista                 6114 non-null   object
dtypes: int64(7), object(15)
memory usage: 1.0+ MB
```

```
1 aeronave.info()
2 aeronave.head()

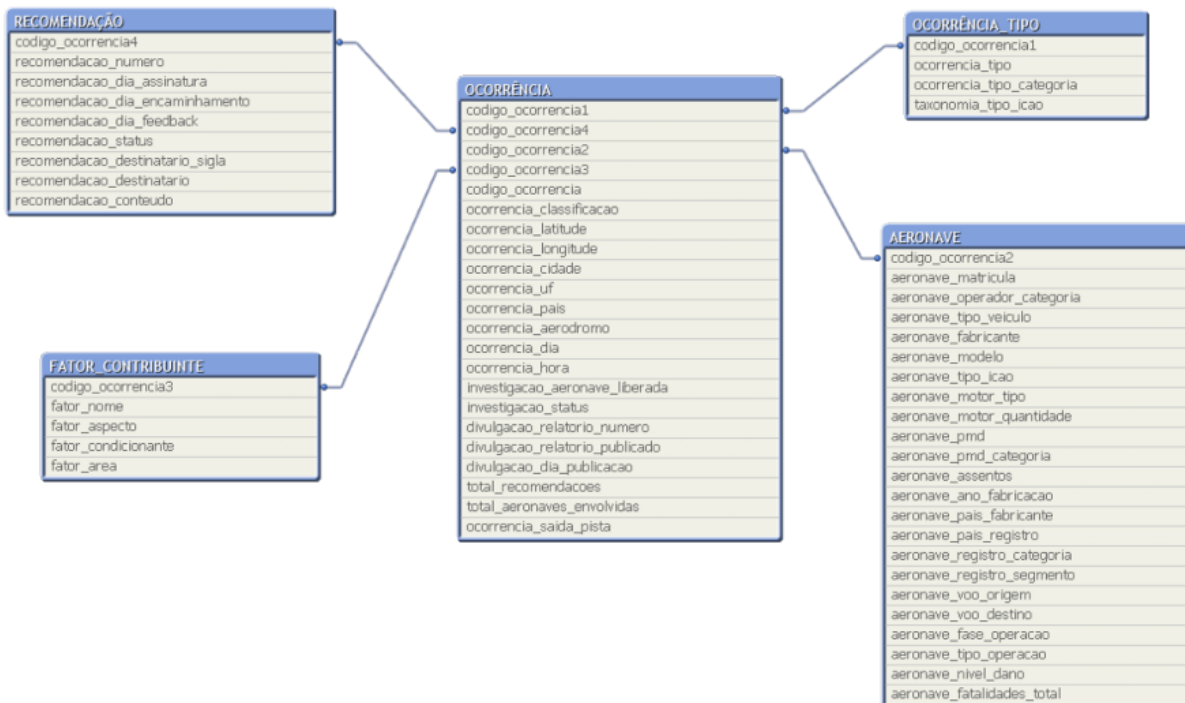
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 6188 entries, 0 to 6187
Data columns (total 23 columns):
#   Column                                Non-Null Count  Dtype
---  -
0   codigo_ocorrencia2                    6188 non-null   int64
1   aeronave_matricula                     6188 non-null   object
2   aeronave_operador_categoria            6188 non-null   object
3   aeronave_tipo_veiculo                  6188 non-null   object
4   aeronave_fabricante                    6188 non-null   object
5   aeronave_modelo                        6188 non-null   object
6   aeronave_tipo_icao                      6188 non-null   object
7   aeronave_motor_tipo                    6185 non-null   object
8   aeronave_motor_quantidade              6188 non-null   object
9   aeronave_pmd                           6188 non-null   int64
10  aeronave_pmd_categoria                  6188 non-null   int64
11  aeronave_assentos                       6010 non-null   float64
12  aeronave_ano_fabricacao                 6005 non-null   float64
13  aeronave_pais_fabricante                6188 non-null   object
14  aeronave_pais_registro                  6188 non-null   object
15  aeronave_registro_categoria            6188 non-null   object
16  aeronave_registro_segmento              6188 non-null   object
17  aeronave_voo_origem                     6187 non-null   object
18  aeronave_voo_destino                     6187 non-null   object
19  aeronave_fase_operacao                  6188 non-null   object
20  aeronave_tipo_operacao                  6188 non-null   object
21  aeronave_nivel_dano                     6188 non-null   object
22  aeronave_fatalidades_total              6188 non-null   int64
dtypes: float64(2), int64(4), object(17)
memory usage: 1.1+ MB
```

```
1 recomendacao_de_seguranca.info()
2 recomendacao_de_seguranca.head()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1872 entries, 0 to 1871
Data columns (total 9 columns):
#   Column                                Non-Null Count  Dtype
---  -
0   codigo_ocorrencia4                    1872 non-null   int64
1   recomendacao_numero                    1872 non-null   object
2   recomendacao_diaassinatura              1872 non-null   object
3   recomendacao_dia_encaminhamento        1872 non-null   object
4   recomendacao_dia_feedback               1146 non-null   object
5   recomendacao_conteudo                   1871 non-null   object
6   recomendacao_status                     1872 non-null   object
7   recomendacao_destinatario_sigla         1872 non-null   object
8   recomendacao_destinatario               1872 non-null   object
dtypes: int64(1), object(8)
memory usage: 131.8+ KB
```

2.3.3) Modelo de dados

Modelo de Dados





TESTE TÉCNICO – CIÊNTISTA DE DADOS A3DATA

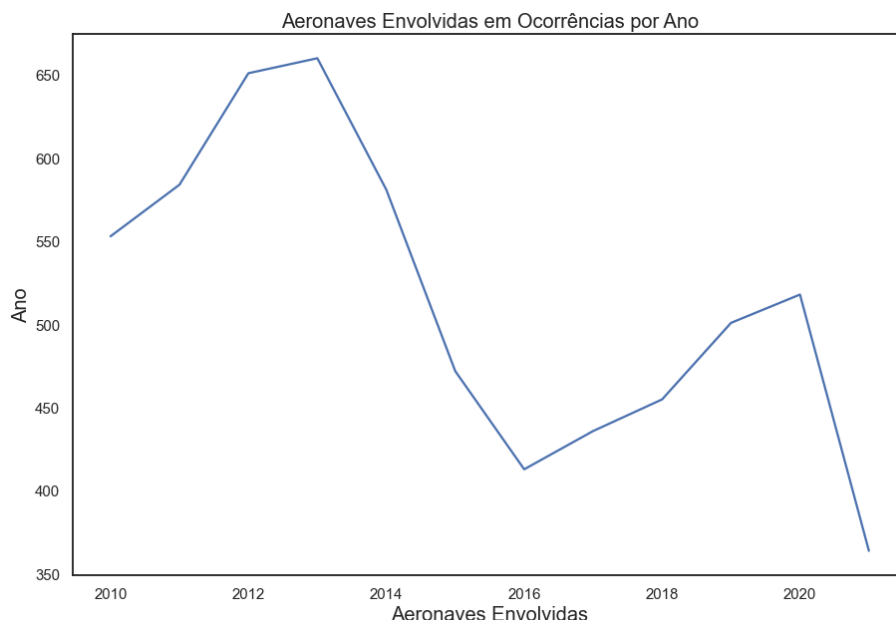
ANTÔNIO DOS SANTOS RAMOS NETO

3) ANÁLISE EXPLORATÓRIA

3.1) Funções

```
1 #Gráficos
2
3 def linha(titulo, eixo_x, eixo_y, data, x, y):
4     sns.set_theme(style="white", context="talk")
5     plt.figure(figsize=(15,10))
6     plt.xticks(fontsize = 15)
7     plt.title(titulo, fontsize = 20)
8     g1 = sns.lineplot(data=data, x=x,y=y)
9     g1.set_xlabel(eixo_x, fontsize = 20)
10    g1.set_ylabel(eixo_y, fontsize = 20)
11    plt.show()
12
13
14 def barra(titulo, eixo_x, eixo_y, data, x, y):
15     sns.set_theme(style="white", context="talk")
16     plt.figure(figsize=(15,10))
17     plt.xticks(fontsize = 15)
18     plt.yticks(fontsize = 15)
19     plt.title(titulo, fontsize = 20)
20    g2 = sns.barplot(data=data, x=x,y=y, palette="dark:slmon_r")
21    g2.set_xlabel(eixo_x, fontsize = 20)
22    g2.set_ylabel(eixo_y, fontsize = 20)
23    plt.show()
24
25 def barra2(titulo1, titulo2, eixo_x, eixo_y, data, x, y, data2, x2, y2):
26     sns.color_palette("rocket")
27     plt.figure(figsize=(16,10))
28
29     plt.subplot(2,2,1)
30     g2 = sns.barplot(data=data, x=x,y=y, palette="ch:s--2,r-.6")
31     plt.title(titulo1, fontsize = 12)
32     g2.set_xlabel(eixo_x, fontsize = 12)
33     g2.set_ylabel(eixo_y, fontsize = 12)
34     plt.xticks(rotation=70, fontsize = 10)
35     plt.yticks(fontsize = 12)
36
37     plt.subplot(2,2,2)
38     g8 = sns.barplot(data=data2, x=x2,y=y2, palette="icefire")
39     plt.title(titulo2, fontsize = 12)
40     g8.set_xlabel(eixo_x, fontsize = 12)
41     g8.set_ylabel(eixo_y, fontsize = 12)
42     plt.xticks(rotation=70, fontsize = 10)
43     plt.yticks(fontsize = 12)
44
45     plt.show()
46
47 def barra2invert(titulo1, titulo2, eixo_x, eixo_y, data, x, y, data2, x2, y2):
48     plt.figure(figsize=(16,15))
49
50     plt.subplot(2,2,1)
51     g2 = sns.barplot(data=data, x=x,y=y, palette="ch:s--2,r-.6")
52     plt.title(titulo1, fontsize = 12)
53     g2.set_xlabel(eixo_x, fontsize = 12)
54     g2.set_ylabel(eixo_y, fontsize = 12)
55     plt.xticks(rotation=70, fontsize = 10)
56     plt.yticks(fontsize = 12)
57
58     plt.subplot(2,2,2)
59     g8 = sns.barplot(data=data2, x=x2,y=y2, palette="icefire")
60     plt.title(titulo2, fontsize = 12)
61     g8.set_xlabel(eixo_x, fontsize = 12)
62     g8.set_ylabel(eixo_y, fontsize = 12)
63     plt.xticks(rotation=70, fontsize = 10)
64     plt.yticks(fontsize = 12)
65
66     plt.show()
67
68
69 def barra2invert3(titulo1, titulo2, eixo_x, eixo_y, eixo_y2, data, x, y, data2, x2, y2, data3, x3, y3, data4, x4, y4):
70     plt.figure(figsize=(33,33))
71
72     plt.subplot(2,2,1)
73     g1 = sns.barplot(data=data, x=x,y=y, palette="ch:s--2,r-.6")
74     plt.title(titulo1, fontsize = 30)
75     g1.set_ylabel(eixo_y, fontsize = 22)
76     plt.xticks(fontsize = 25)
77     plt.yticks(fontsize = 13.5)
78
79     plt.subplot(2,2,2)
80     g2 = sns.barplot(data=data2, x=x2,y=y2, palette="icefire")
81     plt.title(titulo2, fontsize = 30)
82     plt.xticks(fontsize = 25)
83     plt.yticks(fontsize = 13.5)
84
85     plt.subplot(2,2,3)
86     g3 = sns.barplot(data=data3, x=x3,y=y3, palette="ch:s--2,r-.6")
87     g3.set_ylabel(eixo_y2, fontsize = 22)
88     g3.set_xlabel(eixo_x, fontsize = 22)
89     plt.xticks(fontsize = 25)
90     plt.yticks(fontsize = 13.5)
91
92     plt.subplot(2,2,4)
93     g4 = sns.barplot(data=data4, x=x4,y=y4, palette="icefire")
94     g4.set_xlabel(eixo_x, fontsize = 22)
95     plt.xticks(fontsize = 25)
96     plt.yticks(fontsize = 13.5)
97
98 def barra2invert4(titulo1, titulo2, eixo_y, eixo_y2, eixo_y3, eixo_y4, eixo_x, data, x, y, data2, x2, y2, data3, x3, y3, dat
99     plt.figure(figsize=(33,33))
100
101     plt.subplot(4,2,1)
102     g1 = sns.barplot(data=data, x=x,y=y, palette="ch:s--2,r-.6")
103     plt.title(titulo1, fontsize = 30)
104     g1.set_ylabel(eixo_y, fontsize = 22)
105     plt.xticks(fontsize = 25)
106     plt.yticks(fontsize = 13.5)
107
108     plt.subplot(4,2,2)
109     g2 = sns.barplot(data=data2, x=x2,y=y2, palette="icefire")
110     plt.title(titulo2, fontsize = 30)
111     plt.xticks(fontsize = 25)
112     plt.yticks(fontsize = 13.5)
113
114     plt.subplot(4,2,3)
115     g3 = sns.barplot(data=data3, x=x3,y=y3, palette="ch:s--2,r-.6")
116     g3.set_ylabel(eixo_y2, fontsize = 22)
117     plt.xticks(fontsize = 25)
118     plt.yticks(fontsize = 13.5)
119
120     plt.subplot(4,2,4)
121     g4 = sns.barplot(data=data4, x=x4,y=y4, palette="icefire")
122     plt.xticks(fontsize = 25)
123     plt.yticks(fontsize = 13.5)
124
125     plt.subplot(4,2,5)
126     g5 = sns.barplot(data=data5, x=x5,y=y5, palette="ch:s--2,r-.6")
127     g5.set_ylabel(eixo_y3, fontsize = 22)
128     plt.xticks(fontsize = 25)
129     plt.yticks(fontsize = 13.5)
130
131     plt.subplot(4,2,6)
132     g6 = sns.barplot(data=data6, x=x6,y=y6, palette="icefire")
133     plt.xticks(fontsize = 25)
134     plt.yticks(fontsize = 13.5)
135
136     plt.subplot(4,2,7)
137     g7 = sns.barplot(data=data7, x=x7,y=y7, palette="ch:s--2,r-.6")
138     g7.set_ylabel(eixo_y4, fontsize = 22)
139     plt.xticks(fontsize = 25)
140     plt.yticks(fontsize = 13.5)
141
142     plt.subplot(4,2,8)
143     g8 = sns.barplot(data=data8, x=x8,y=y8, palette="icefire")
144     g8.set_xlabel(eixo_x, fontsize = 22)
145     plt.xticks(fontsize = 25)
146     plt.yticks(fontsize = 13.5)
147
148     plt.show()
149
150
151 def box2(titulo, titulo2, eixo_x, eixo_y, data, x, y, data2, x2, y2):
152     plt.figure(figsize=(15,10))
153
154     plt.subplot(2,2,1)
155     sns.set_theme(style="white", context="talk")
156     plt.xticks(rotation=70, fontsize = 10)
157     plt.yticks(fontsize = 15)
158     plt.title(titulo, fontsize = 13)
159     g3 = sns.boxplot(data=data, x=x,y=y, palette="ch:s--2,r-.6")
160     g3.set_xlabel(eixo_x, fontsize = 13)
161     g3.set_ylabel(eixo_y, fontsize = 13)
162
163     plt.subplot(2,2,2)
164     sns.set_theme(style="white", context="talk")
165     plt.xticks(rotation=70, fontsize = 10)
166     plt.yticks(fontsize = 15)
167     plt.title(titulo2, fontsize = 13)
168     g4 = sns.boxplot(data=data2, x=x2,y=y2, palette="icefire")
169     g4.set_xlabel(eixo_x, fontsize = 13)
170     g4.set_ylabel(eixo_y, fontsize = 13)
171
172     plt.show()
173
174 def histograma2(titulo, titulo2, eixo_x, eixo_y, data, x, data2, x2):
175     plt.figure(figsize=(15,10))
176
177     plt.subplot(2,2,1)
178     sns.set_theme(style="white", context="talk")
179     plt.xticks(rotation=0, fontsize = 15)
180     plt.yticks(fontsize = 15)
181     plt.title(titulo, fontsize = 15)
182     g5 = sns.histplot(data, x=x, state='count', palette="ch:s-.25,r-.25")
183     g5.set_xlabel(eixo_x, fontsize = 15)
184     g5.set_ylabel(eixo_y, fontsize = 15)
185
186     plt.subplot(2,2,2)
187     sns.set_theme(style="white", context="talk")
188     plt.xticks(rotation=0, fontsize = 15)
189     plt.yticks(fontsize = 15)
190     plt.title(titulo2, fontsize = 15)
191     g6 = sns.histplot(data2, x=x2, state='count', palette="fiare")
192     g6.set_xlabel(eixo_x, fontsize = 15)
193     g6.set_ylabel(eixo_y, fontsize = 15)
194
195     plt.show()
196
197 #Conversões e tratamentos de dados
198
199 def formatar_eixos (coluna, nomecoluna, nomeanterior):
200     tipo_de_ocorrendia = pd.DataFrame(coluna.value_counts())
201     tipo_de_ocorrendia = tipo_de_ocorrendia.reset_index().head(10)
202     tipo_de_ocorrendia = tipo_de_ocorrendia.rename(columns={"index": nomecoluna, nomeanterior: "quantidade"})
203     return tipo_de_ocorrendia
204
205 def encarta_eixo_x (eixo,novovalor):
206     novovalor = []
207     for n in range(len(eixo)):
208         a = eixo[n][14] + "..."
209         novovalor.append(a)
210     return novovalor
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
```

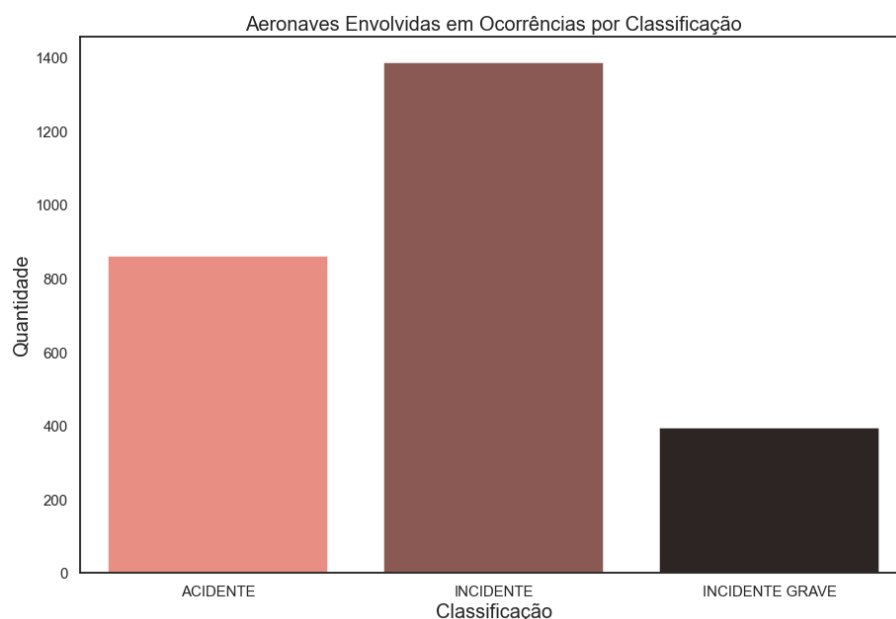
3.2) Investigação



A base vem sendo atualizada desde 2010 e na análise gráfica acima podemos perceber que houve uma redução significativa do aeronaves envolvidas em ocorrências do ano de 2012 até 2016, porém de 2016 até 2020 esse número vem novamente tendo um crescimento considerável, o que demonstra que há necessidade do estudo desses dados para pensar em medidas preventivas para segurança da aviação civil brasileira.

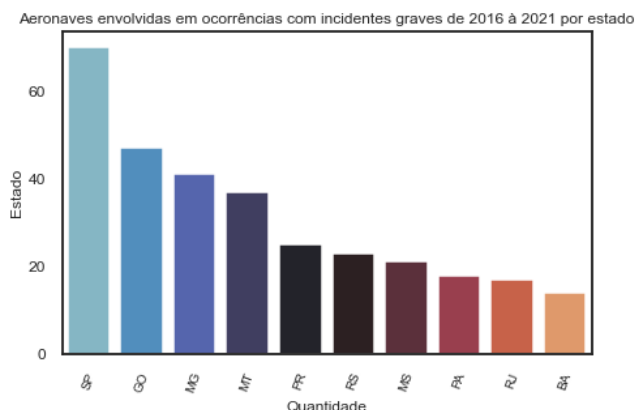
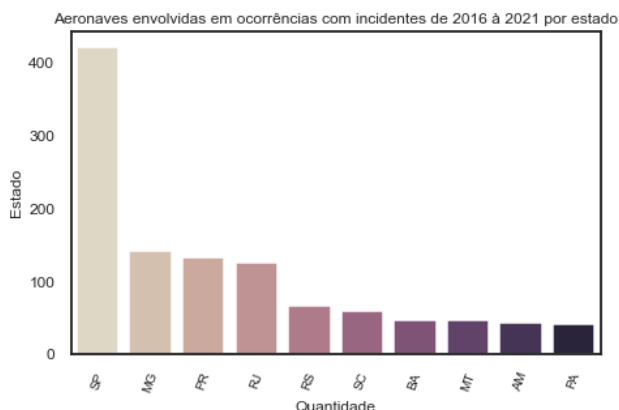
3.2.1) Identificação dos maiores eventos de ocorrência

3.2.1.1) Investigando por Classificação



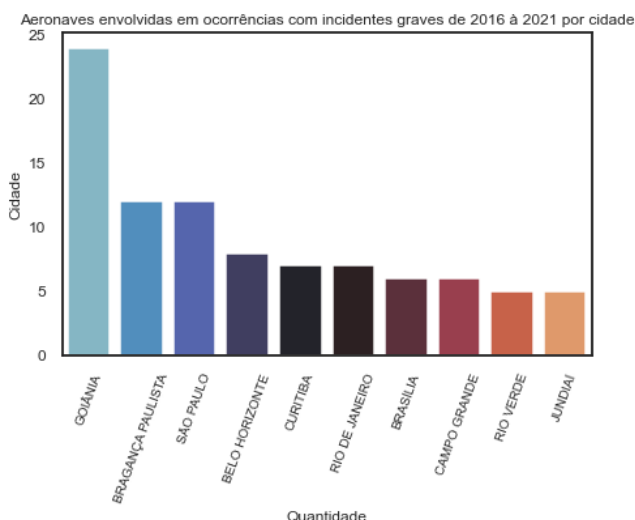
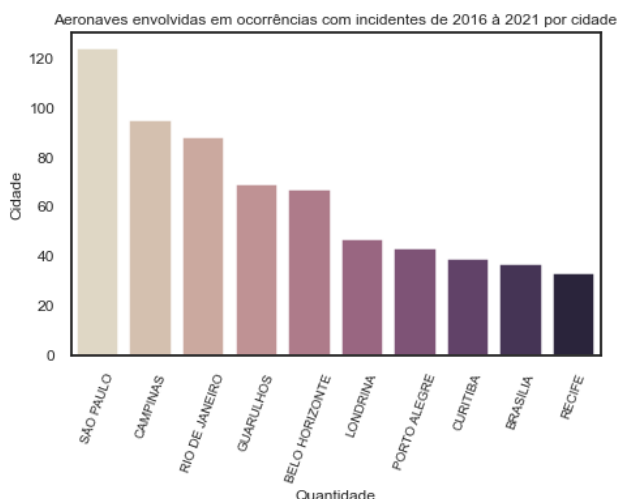
É possível identificar que a classificação com o maior número de ocorrências faz parte do tipo incidente, aproximadamente 52%, porém 15% são de classificação de incidência grave. Ao longo dos últimos 6 anos, temos então uma média de 66 incidentes graves por ano o que é um número relevante. Será investigado mais a fundo os dois tipos de classificação, incidente e incidente grave.

3.2.1.2) Investigando por Estado



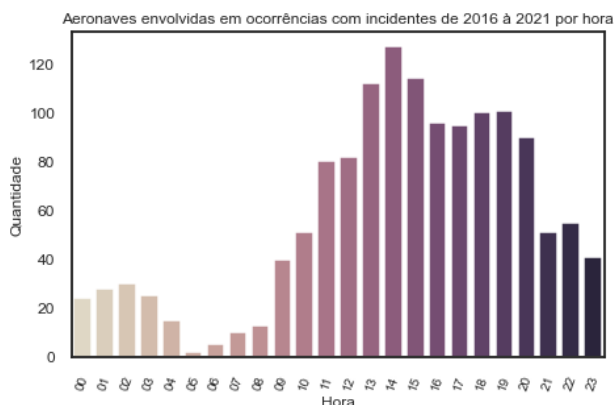
Dos 27 estados brasileiros ao longo de 12 anos (2010 - nov/2021), 10 deles possuem 80% dos incidentes com aeronaves, com ênfase em SP, MG, PR, RJ que representam 59% dos incidentes. Já para os incidentes graves, dos 27 estados brasileiros ao longo de 12 anos (2010 - nov/2021), 10 deles possuem 79% dos incidentes graves, com ênfase também em SP, GO, MG, MT que representam 39% dos incidentes graves.

3.2.1.3) Investigando por Cidade



As 10 principais cidades com a maior ocorrência de incidentes representam 46%. As 10 principais cidades com a maior ocorrência de incidentes graves representam 23%, é possível também observar que embora Goiás esteja em terceiro lugar em incidentes graves, a cidade de Goiânia é a que mais teve incidentes graves ao longo do período.

3.2.1.4) Investigando por hora



É possível identificar que os picos mais frequentes nas ocorrências com incidentes são no intervalo das 13-15 horas. É possível identificar que os picos mais frequentes nas ocorrências com incidentes graves acontece às 12 horas e no intervalo de 19-20 horas.

Resumo

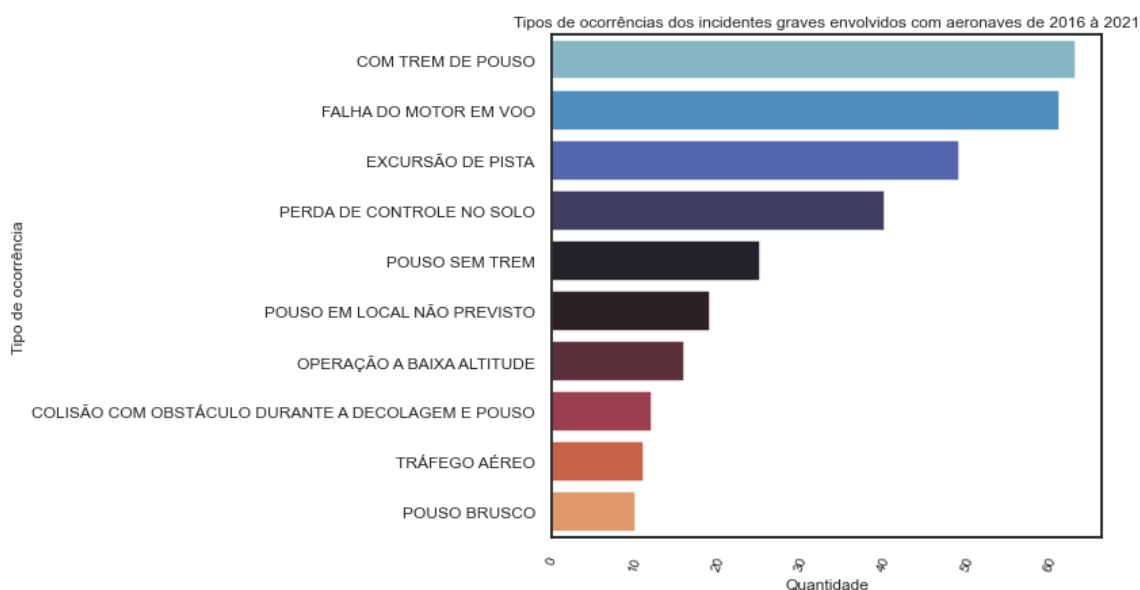
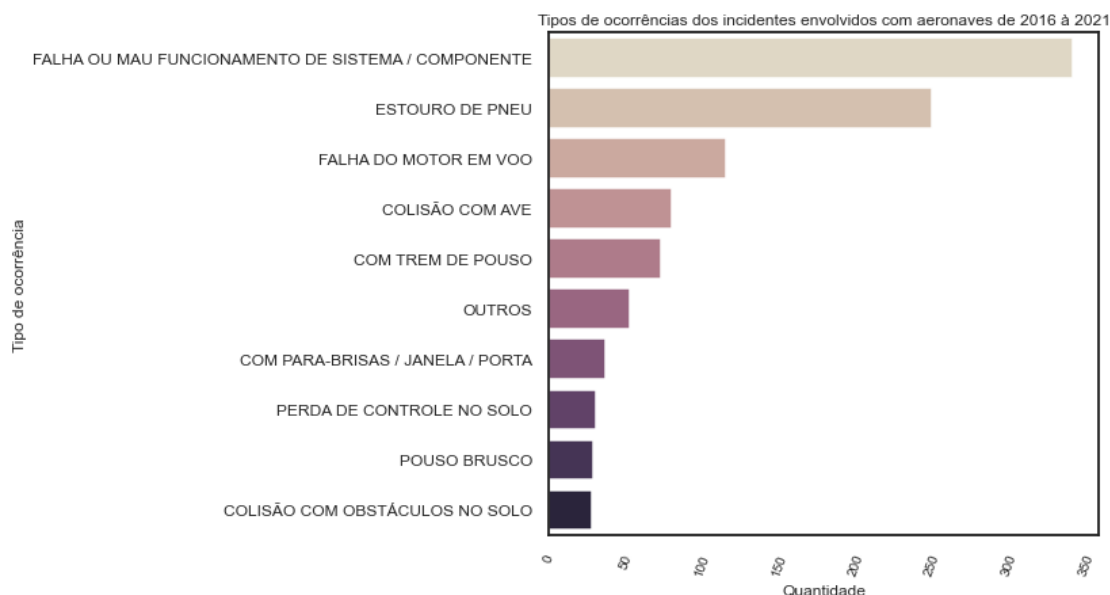
Portanto, no momento de Identificação dos maiores eventos de ocorrência conseguimos definir alguns pontos de partida para investigação maior dos tipos de ocorrências, fator de contribuições e aeronaves, eles são:

Ocorrências com incidentes: Dos 27 estados brasileiros ao longo de 12 anos (2010 - nov/2021), 10 deles possuem 80% dos incidentes com aeronaves, com ênfase em SP, MG, PR, RJ que representam 59% dos incidentes. As 10 principais cidades com a maior ocorrência de incidentes representam 46%. É possível identificar que os picos mais frequentes nas ocorrências com incidentes são no intervalo das 13-15 horas.

Ocorrências com incidentes graves: Já para os incidentes graves, dos 27 estados brasileiros ao longo de 12 anos (2010 - nov/2021), 10 deles possuem 79% dos incidentes graves, com ênfase também em SP, GO, MG, MT que representam 39% dos incidentes graves. As 10 principais cidades com a maior ocorrência de incidentes graves representam 23%, é possível também observar que embora Goiás esteja em terceiro lugar em incidentes graves, a cidade de Goiânia é a que mais teve incidentes graves ao longo do período. É possível identificar que os picos mais frequentes nas ocorrências com incidentes graves acontece às 12 horas e no intervalo de 19-20 horas.

3.2.2) Investigando o tipo de ocorrência, fator de contribuição e aeronave

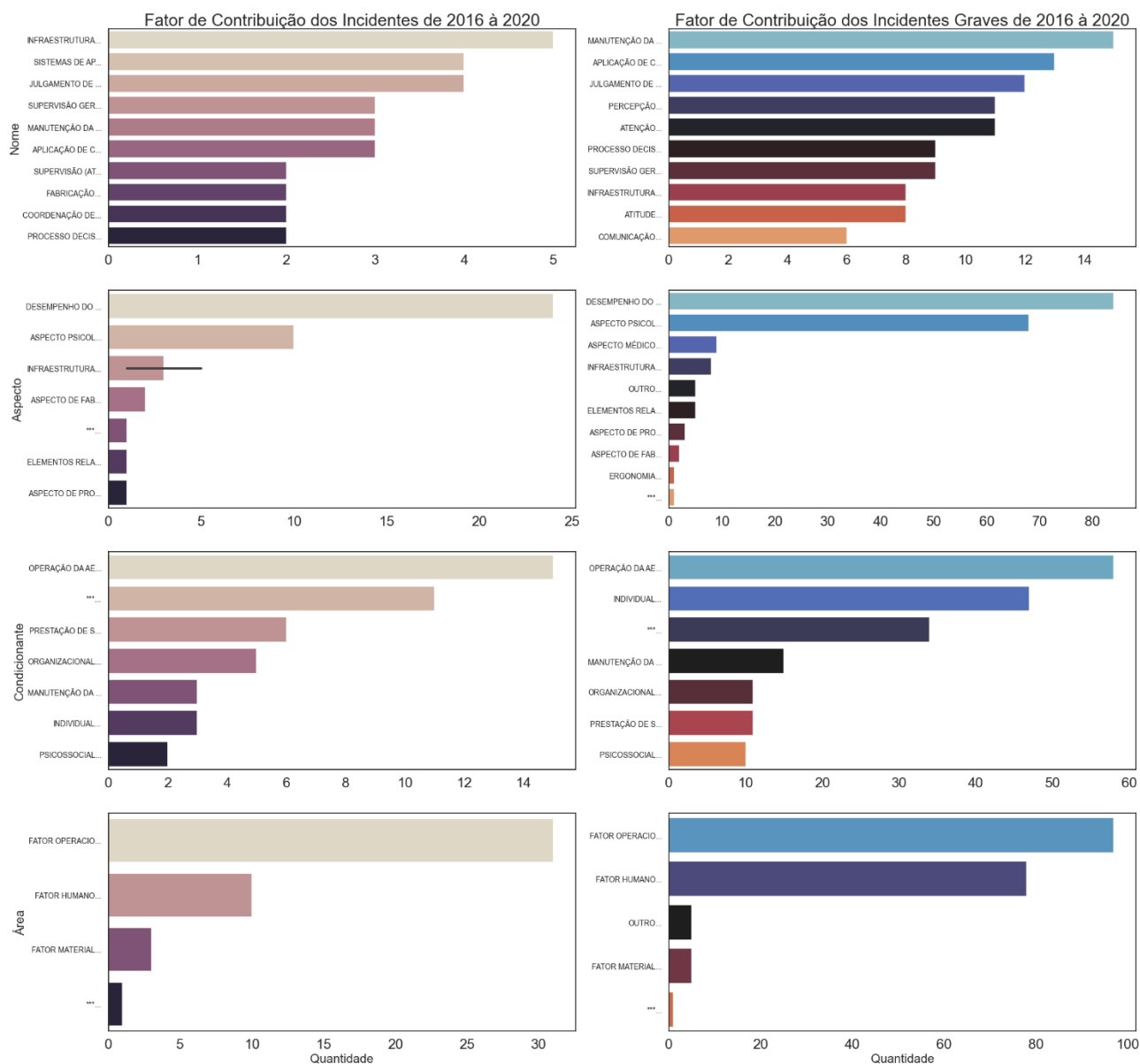
3.2.2.1) Investigando o tipo de ocorrência



Dos 10 tipos de ocorrências mais frequentes para incidentes, os 3 que tiveram maior recorrência de 2016 à 2021 representam 50%, são eles: 1) Falha ou mau funcionamento de sistema, 2) Estouro de pneu, 3) Falha do motor em voo. Já dos 10 tipos de ocorrências mais frequentes em incidentes graves temos a presença de 4 mais recorrentes equivalendo 53%, são elas: 1) Com trem de pouso, 2) Falha do motor em voo, 3) excursão de pista.

É possível perceber que há uma relevante frequência de ocorrências tanto em incidentes leves como nos graves de falha do motor e falha de sistemas de trem de pouso ou outros.

3.2.2.2) Investigando o fator de contribuição



Acima investigamos os tipos de ocorrências e agora podemos identificar os fatores que contribuíram para esses tipos de ocorrências acontecerem:

Fator de contribuição para os incidentes: Dos três fatores em incidentes mais frequentes citados acima (1- Falha ou mau funcionamento de sistema, 2- Estouro de pneu, 3- Falha do motor em voo) temos os seus respectivos fatores de contribuições mais relevantes:

- 1) Nome - Infraestrutura Aeroportuária / Sistemas De Apoio / Julgamento De Pilotagem
- 2) Aspecto - Desempenho Do Ser Humano / Aspecto Psicológico
- 3) Condicionante - Operação Da Aeronave

4) Área - Fator Operacional

É possível perceber que mesmo o problema tendo sido operacional, muito esteve envolvido o desempenho humano nesse tipo de problema, que por mais que os incidentes leves tenham acontecido por falhas na aeronave, as manutenções, checagem de pneus, entre outros são feitos por pessoas e esses procedimentos precisam ser bem checados.

Fator de contribuição para os incidentes graves: Dos três fatores em incidentes mais frequentes citados acima (1- Com trem de pouso, 2- Falha do motor em voo, 3- excursão de pista) temos os seus respectivos fatores de contribuições mais relevantes:

1) Nome - Manutenção Da Aeronave / Aplicação De Comandos / Julgamento De Pilotagem

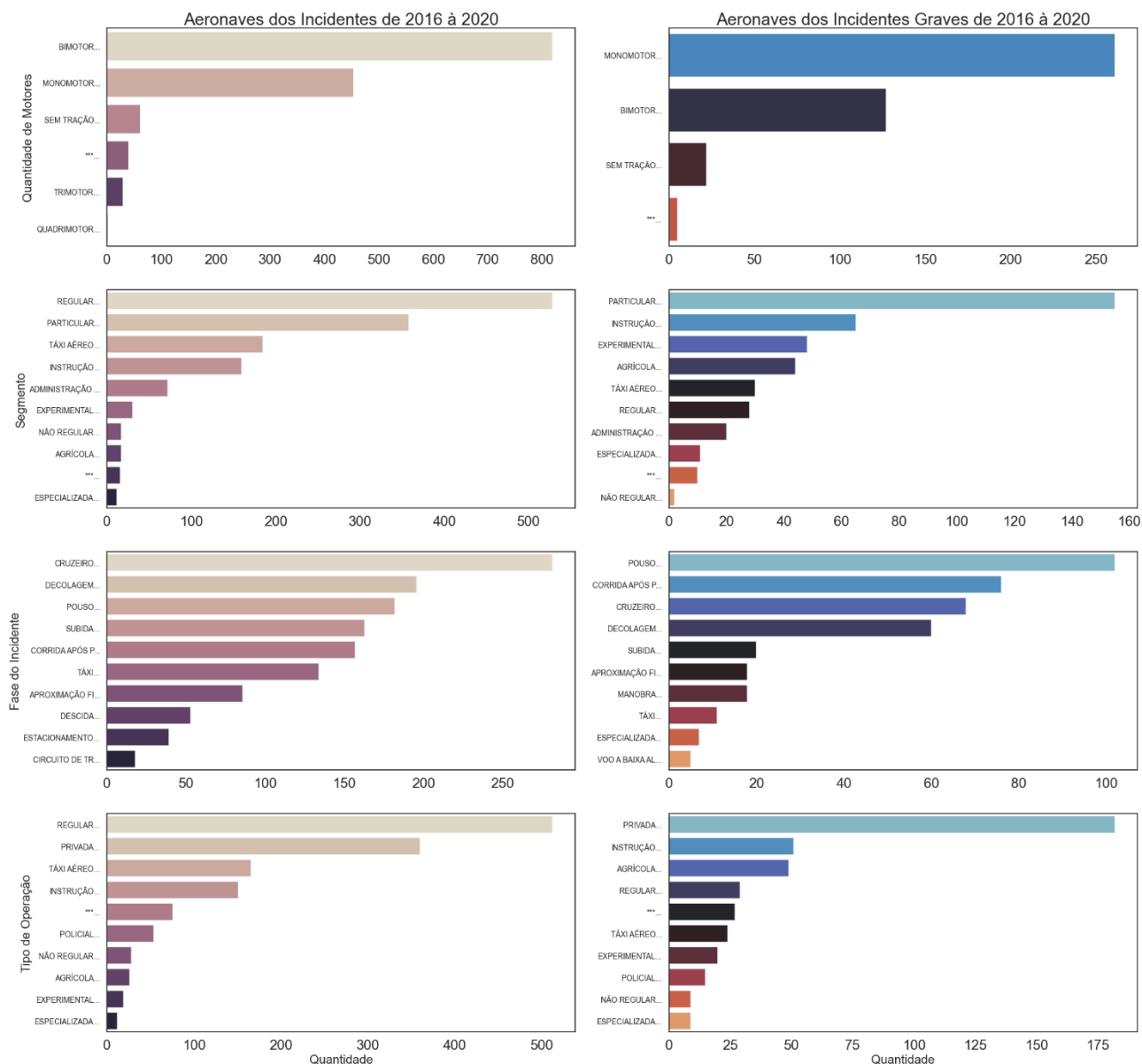
2) Aspecto - Desempenho Do Ser Humano / Aspecto Psicológico

3) Condicionante - Operação Da Aeronave / Individual

4) Área - Fator Operacional / Fator Humano

Nos incidentes fatais o impacto das manutenções é um dos fatores mais relevantes, porém o fator humano também continua sendo aspecto recorrente, demonstrando ainda mais que treinamentos, procedimentos precisam ser reformulados para evitar os incidentes graves.

3.2.2.3) Investigando a aeronave



Além de ter estudado os tipos de ocorrências e os fatores gerados, também é importante entender como isso impacta nos diferentes tipos de aeronaves para perceber se os problemas operacionais e humanos podem ser melhorados em específico.

Informações das aeronaves nos incidentes: Dos três principais fatores de contribuições acima (1- Infraestrutura Aeroportuária, 2- Sistemas De Apoio, 3-Julgamento De Pilotagem), as principais informações das aeronaves nos acontecimentos foram:

- 1) Quantidade De Motores - Bimotor / Monomotor
- 2) Segmento - Regular / Particular / Táxi Aéreo
- 3) Fase Do Incidente - Cruzeiro / Decolagem / Pouso
- 4) Tipo De Operação - Regular / Privada

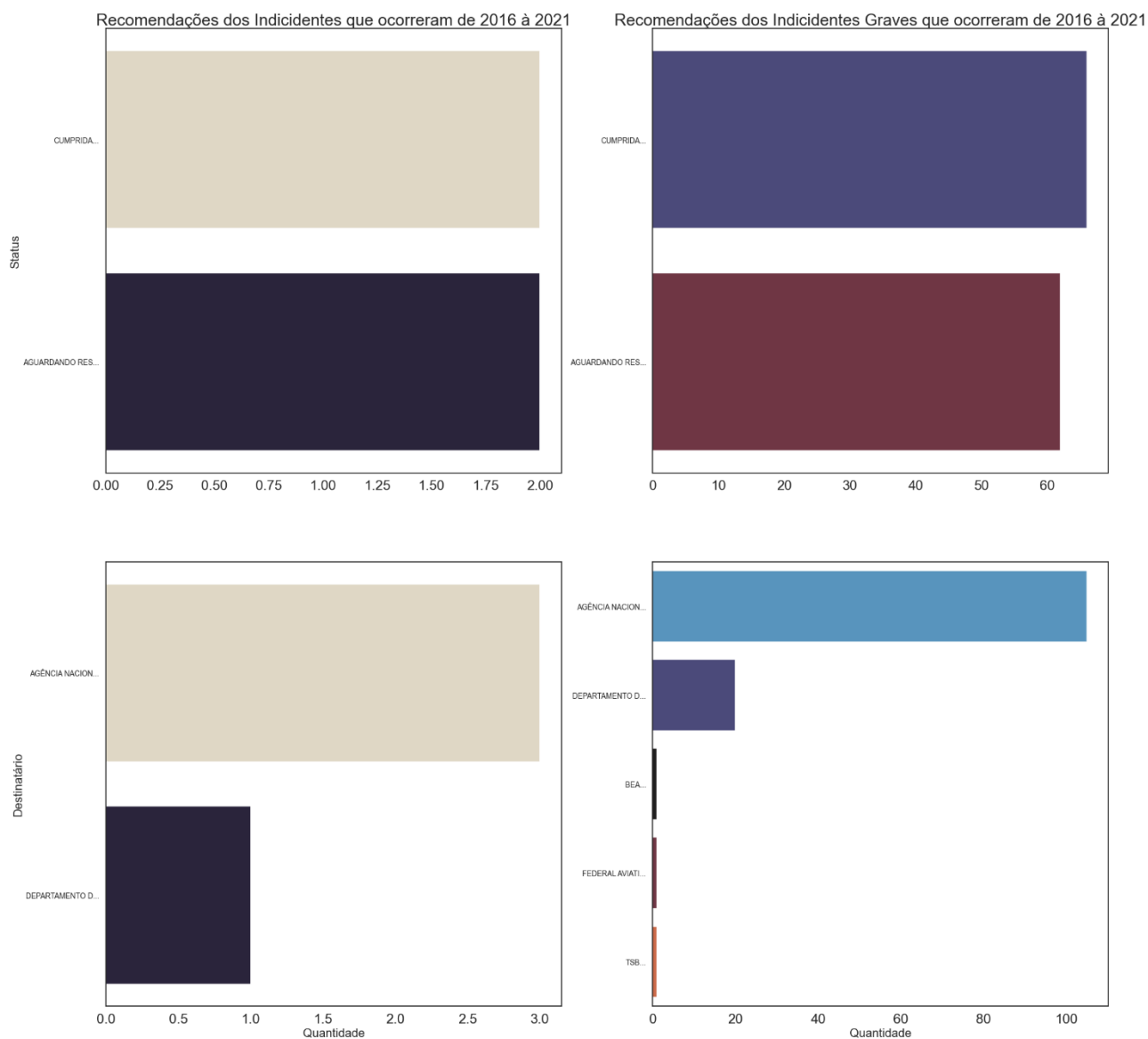
Percebe-se que a maioria dos fatores de contribuições foram em voos pequenos nos casos dos incidentes e o fato de terem acontecido muitos fatores relacionados a sistemas de apoio e infraestrutura aeroportuária pode ser associado aos pequenos voos em aeroportos menores e por consequência com exigências menores de legislação.

Informações das aeronaves nos incidentes graves: Dos três principais fatores de contribuições acima (1-Manutenção Da Aeronave, 2-Aplicação De Comandos, 3-Julgamento De Pilotagem), as principais informações das aeronaves nos acontecimentos foram:

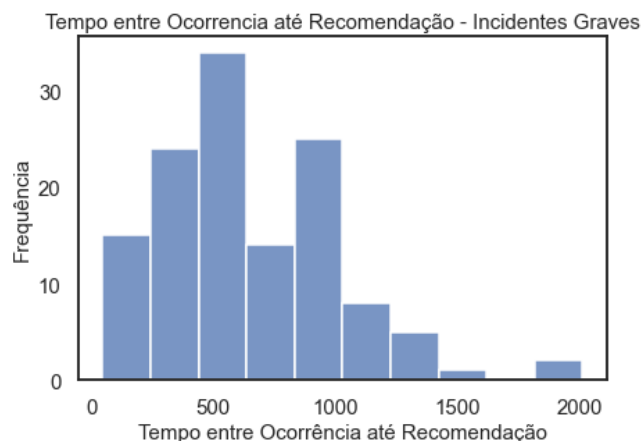
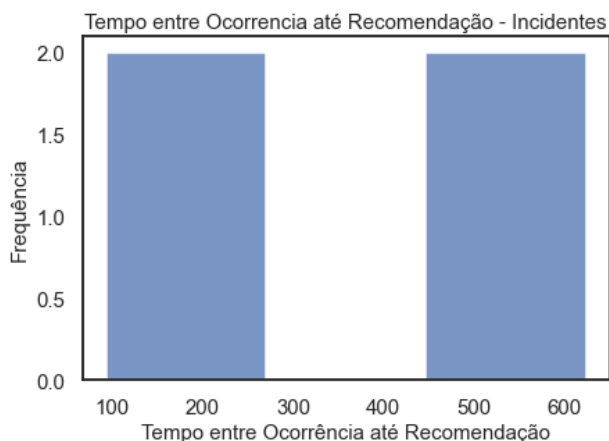
- 1) Quantidade De Motores - Monomotor / Bimotor
- 2) Segmento - Particular / Instrução / Experimental
- 3) Fase Do Incidente - Pouso / Corrida Após Pouso
- 4) Tipo De Operação - Privada / Instrução

Nos voos onde aconteceram incidentes graves, a maioria aconteceu também em voos pequenos o que também pode recorrer a legislação mais leve para esse tipo de operação, além disso, vale ressaltar que a fase desses voos aconteceu quase que em totalidade no pouso o que justifica os fatores de contribuição serem humanos.

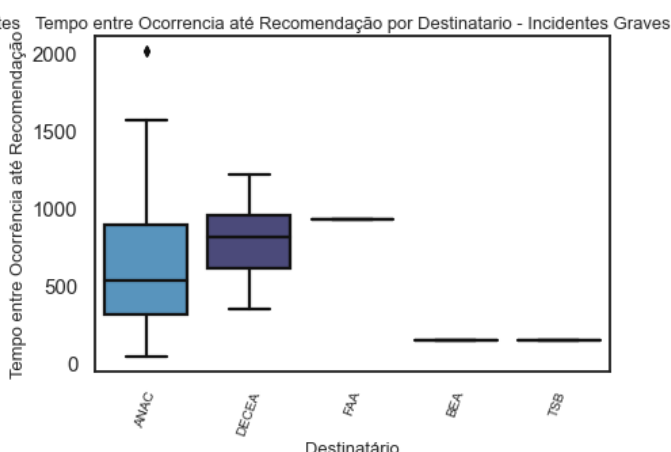
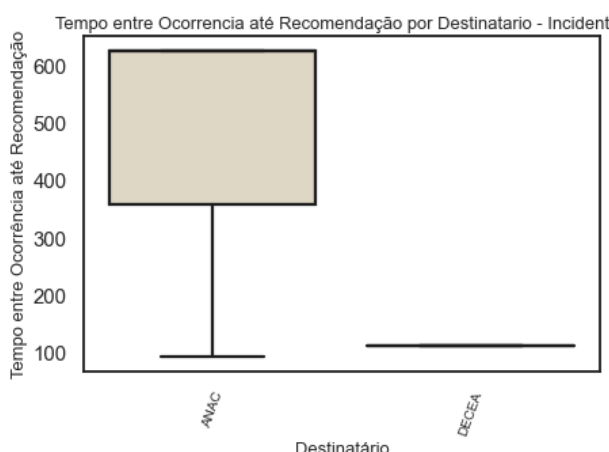
3.2.3) Investigando as recomendações



Embora se tenha poucos dados referentes as recomendações para os incidentes, podemos perceber que tanto para os incidentes leves e graves de 2016 até 2021 temos aproximadamente 50% em andamento e com mais de 90% sendo a Agência Nacional de Aviação Civil sendo responsáveis pela investigação.



Como mencionado acima, embora os poucos dados em relação as recomendações dos incidentes comprometam a análise, no caso dos incidentes graves podemos perceber que há um tempo de recomendação muito longo, o que pode comprometer a reincidência de acidentes por falhas que poderiam ser evitáveis.



Novamente observando que temos poucos dados em relação as recomendações para o caso dos incidentes leves, porém nos casos de incidentes graves onde temos mais tempos podemos perceber que a ANAC que representa 25% das recomendações dos 396 incidentes graves e 82% dos 128 investigações abertas possui uma variabilidade muito alta para liberação da recomendação, fato esse que pode estar acontecendo por conta do acúmulo de ocorrências, uma vez que a DECEA que obtém 15% das recomendações abertas possui uma variabilidade menor. Toda via, a ANAC ainda possui uma mediana menor em relação a DECEA, porém ainda muito alto, tendo casos de mais de 2 anos.

Resumo

Portanto, na investigação em relação as recomendações conseguimos identificar alguns pontos cruciais em relação ao status, destinatário e tempo de recomendação:

Embora se tenha poucos dados referentes as recomendações para os incidentes, podemos perceber que tanto para os incidentes leves e graves de 2016 até 2021 temos aproximadamente 50% em andamento e com mais de 90% sendo a Agência Nacional de Aviação Civil sendo responsáveis pela investigação. Como mencionado acima, embora os poucos dados em relação as recomendações dos incidentes comprometam a análise, no



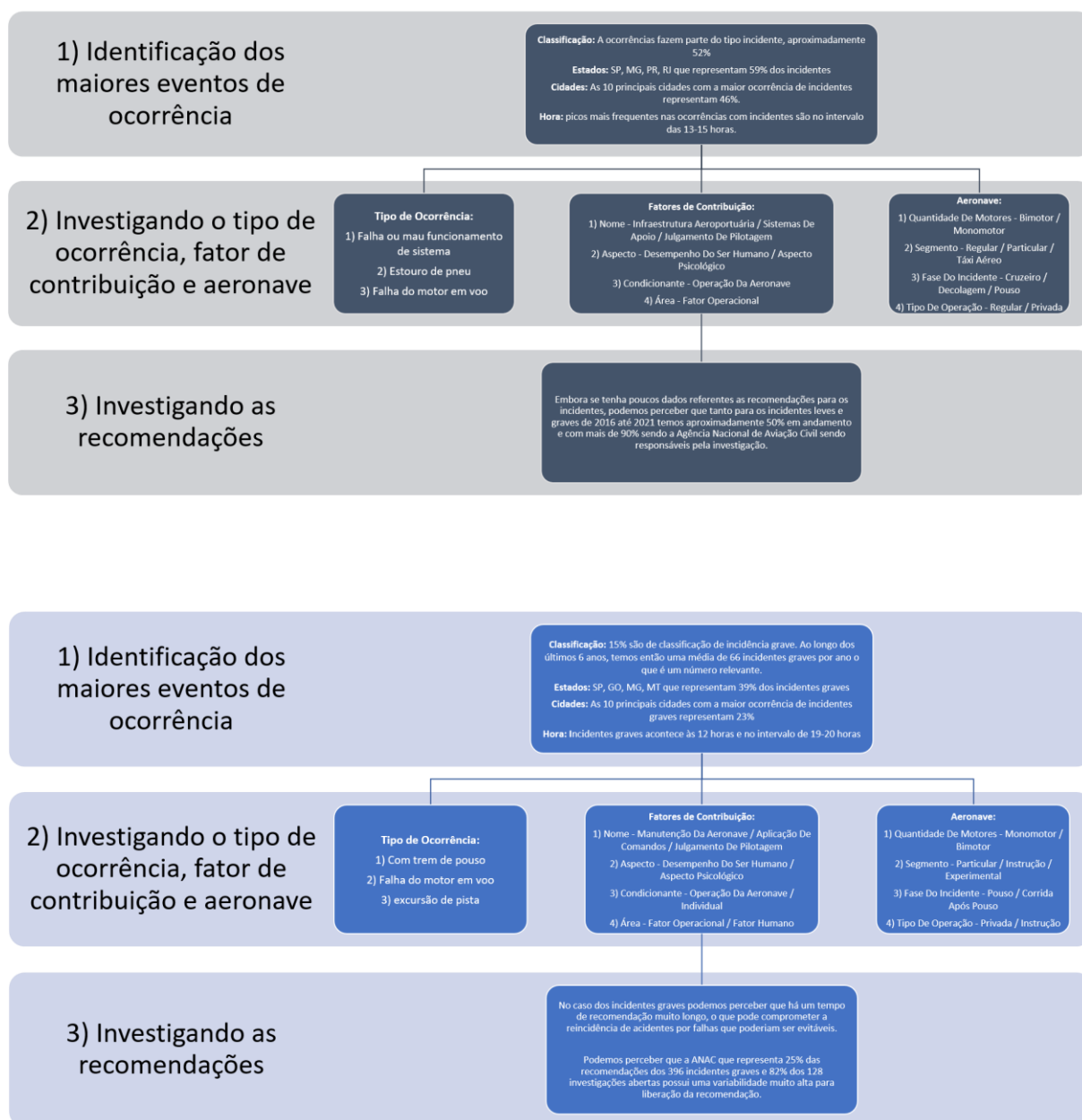
TESTE TÉCNICO – CIÊNTISTA DE DADOS A3DATA

ANTÔNIO DOS SANTOS RAMOS NETO

caso dos incidentes graves podemos perceber que há um tempo de recomendação muito longo, o que pode comprometer a reincidência de acidentes por falhas que poderiam ser evitáveis. Novamente observando que temos poucos dados em relação as recomendações para o caso dos incidentes leves, porém nos casos de incidentes graves onde temos mais tempos podemos perceber que a ANAC que representa 25% das recomendações dos 396 incidentes graves e 82% dos 128 investigações abertas possui uma variabilidade muito alta para liberação da recomendação, fato esse que pode estar acontecendo por conta do acúmulo de ocorrências, uma vez que a DECEA que obtém 15% das recomendações abertas possui uma variabilidade menor. Toda via, a ANAC ainda possui uma mediana menor em relação a DECEA, porém ainda muito alto, tendo casos de mais de 2 anos.

4) HIPÓTESES LEVANTADAS

Com a exploração feita acima, alguns insights surgiram referentes a linha de raciocínio e hipóteses unidas a ações para mitigar possíveis riscos e propor melhorias para aviação, abaixo temos o **resumo dos incidentes e incidentes graves**:



De acordo com a perspectiva seguida, temos as seguintes hipóteses e ações de melhoria propostas:

a) Hipótese: Um dos fatores para a maior recorrência dos incidentes e incidentes graves acontecerem em voos particulares pode estar associado ao fato dos pilotos terem muitas atribuições para operacionalização dos voos. Enquanto nos voos comerciais os pilotos cuidam somente do plano de voo e a pilotagem, os pilotos de voos particulares, táxi

aéreo em sua maioria resolvem muitas coisas da operação, o que pode impactar no desgaste e aumento da probabilidade de ocorrência de incidentes.

Ação: Exigir que as empresas tenham a mínima estrutura para garantir maior segurança para a operacionalização.

b) Hipótese: O tempo de investigação dos incidentes e incidentes graves é muito elevado devido a concentração das investigações pela ANAC o que pode recorrer em mais incidentes pela demora de aplicação das recomendações.

Ação: Coparticipação nas investigações com demais órgãos para agilizar o processo de recomendação.

c) Hipótese: É possível que os incidentes e incidentes graves ocorram em sua maioria nos principais estados por conta do volume dos voos, porém o fato não justifica a grande quantidade de incidentes por falha de sistemas uma vez que os aeródromos possuem uma melhor infraestrutura. A hipótese é que esses voos em sua maioria particulares e táxi aéreos talvez estejam economizando em manutenção e checagem das aeronaves.

Ação: Maior legislação para aumentar a garantia da segurança dos voos.

d) Hipótese: Apesar das ocorrências serem em maioria de riscos operacionais para os incidentes, porém o fator humano ainda é intrínseco ao processo, portanto uma das hipóteses é que por serem voos pequenos, a exigência de uma maior capacitação, horas de voos e experiência não é uma das maiores prioridades para o piloto.

Ação: Exigir maiores capacitações e horas de voos teste para pilotos poderem retirar a carteira de aviação.

5) CONCLUSÕES E INSIGHTS GERADOS

Por fim, ao final do estudo foi possível cumprir com o objetivo de conduzir uma análise exploratória para identificar o comportamento dos dados do CENIPA de 2016 à 2020 que por sua vez obteve um aumento de ocorrência. Além disso, o estudo teve foco em investigar os "incidentes" pelo alto volume e os "incidentes graves" para evitar fatalidades na aviação.

Ao longo o estudo segmentado em 3 investigações foi possível entender um pouco mais sobre as ocorrências, tipo de ocorrência, fator de contribuição, aeronave e recomendações. Com isso, os principais insights geradores das ocorrências de incidentes e incidentes graves foram:

- 1) Excesso de atribuições para os pilotos da aviação particular/táxi aéreo;
- 2) Tempo de aplicação de recomendação longo;
- 3) Baixa investigação da operacionalização dos serviços de voos pequenos;
- 4) Baixa exigência de capacitações e horas de voos.

Além disso, também foram propostas ações para mitigar os riscos acima citados:

- 1) Exigir que as empresas tenham a mínima estrutura para garantir maior segurança para a operacionalização;
- 2) Coparticipação nas investigações com demais órgãos para agilizar o processo de recomendação;
- 3) Maior legislação para aumentar a garantia da segurança dos voos;
- 4) Exigir maiores capacitações e horas de voos teste para pilotos poderem retirar a carteira de aviação.