

Screen term data

Dominik Glandorf

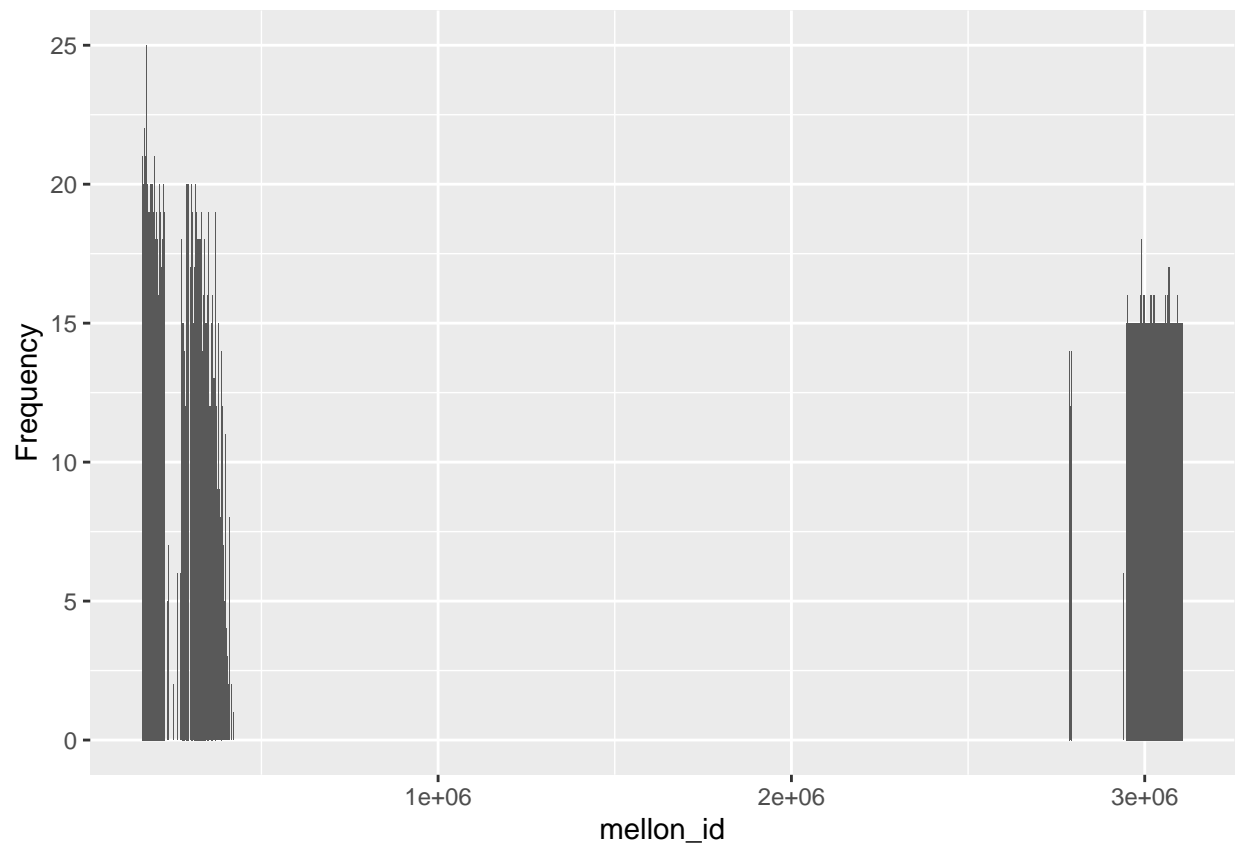
2023-05-11

Student by term

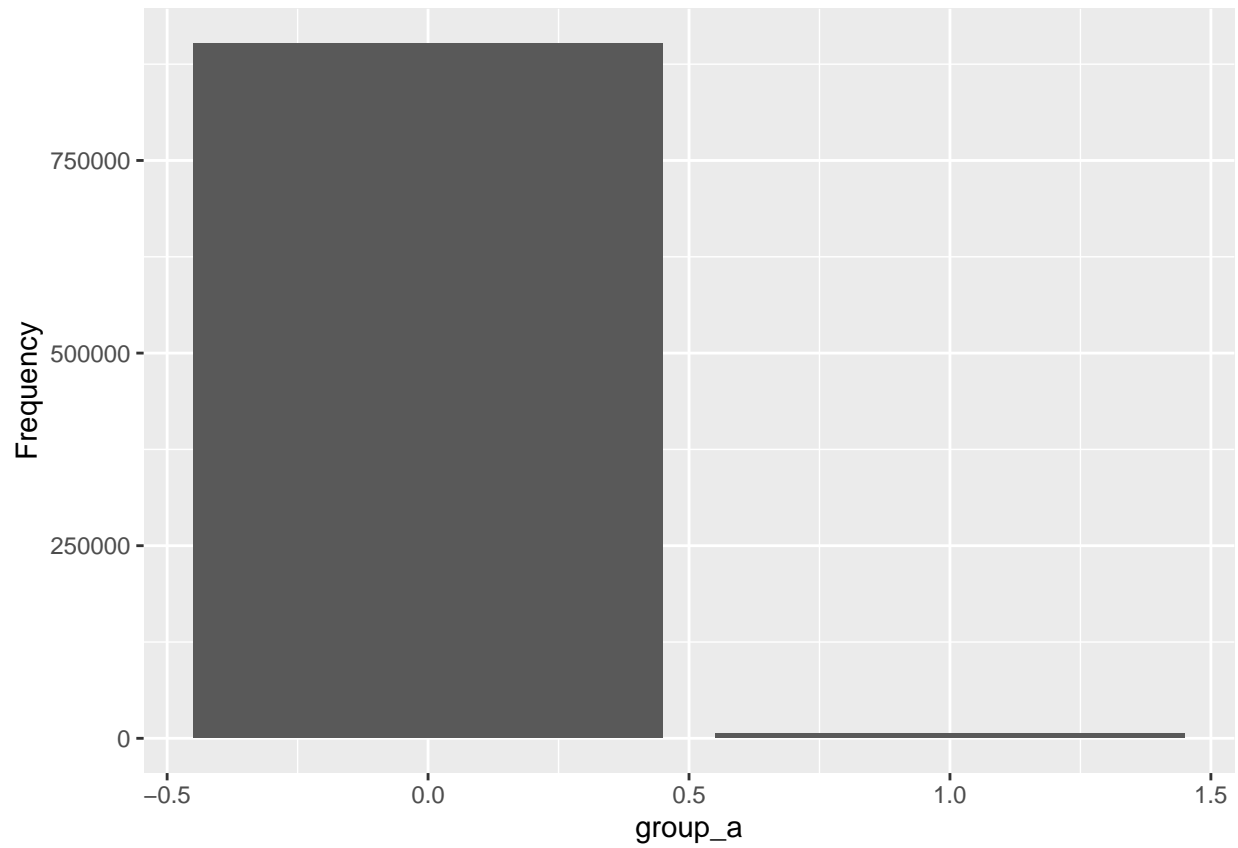
The spreadsheet knows 159 variables. The corresponding file contains 179 columns. The expected variables `student_term_sid (pk)`, `class_standing`, `cumulative_units_attempted_graduate`, `cumulative_units_attempted_online`, `cumulative_units_attempted_transfer` are not in the dataset. Therefore exist variables `created_dttm`, `current_units_attempted_transfer`, `current_units_attempted_online`.

Available variables

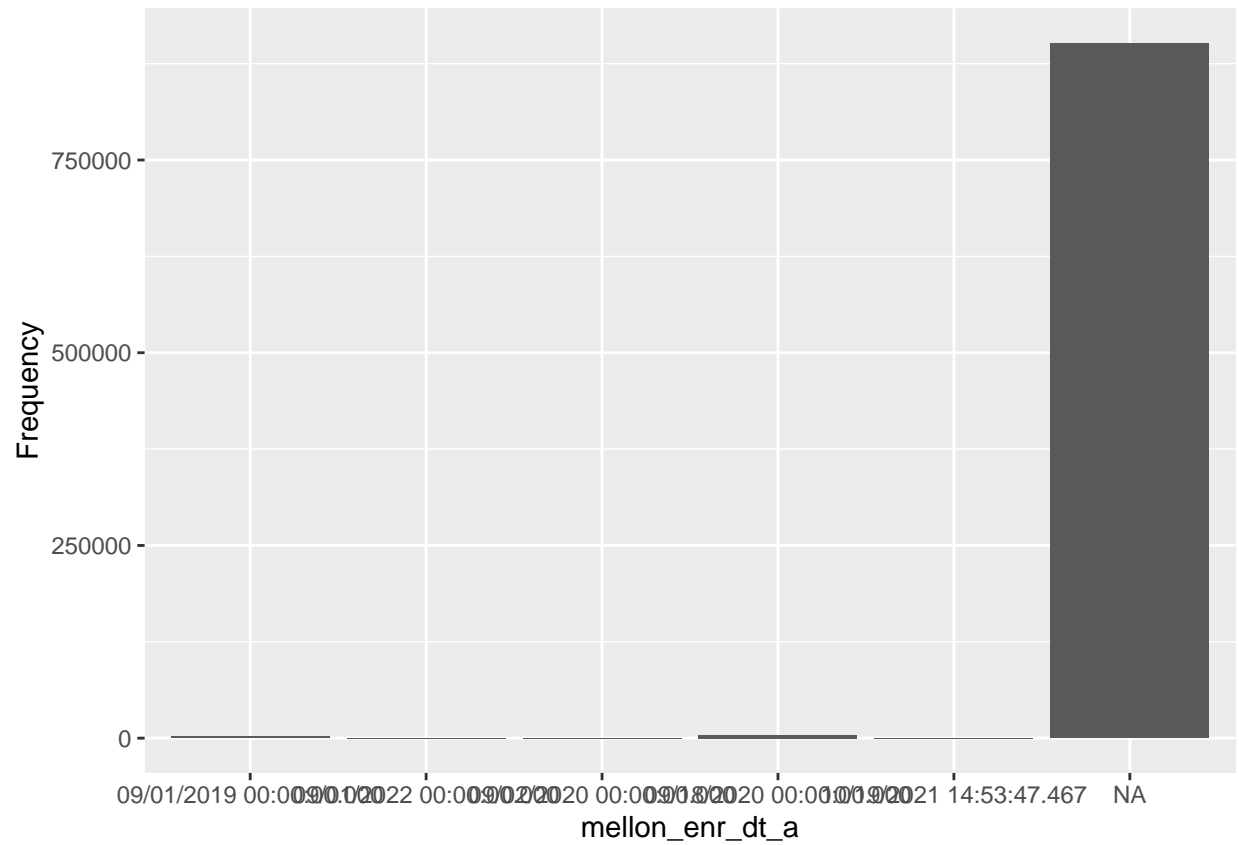
```
[1] Variable: mellon_id, type: numeric, requested: yes
[1] Values (97021 unique): 162767, 162770, 162771, 162774, 162778, ...
[1] Missing: 0%
```



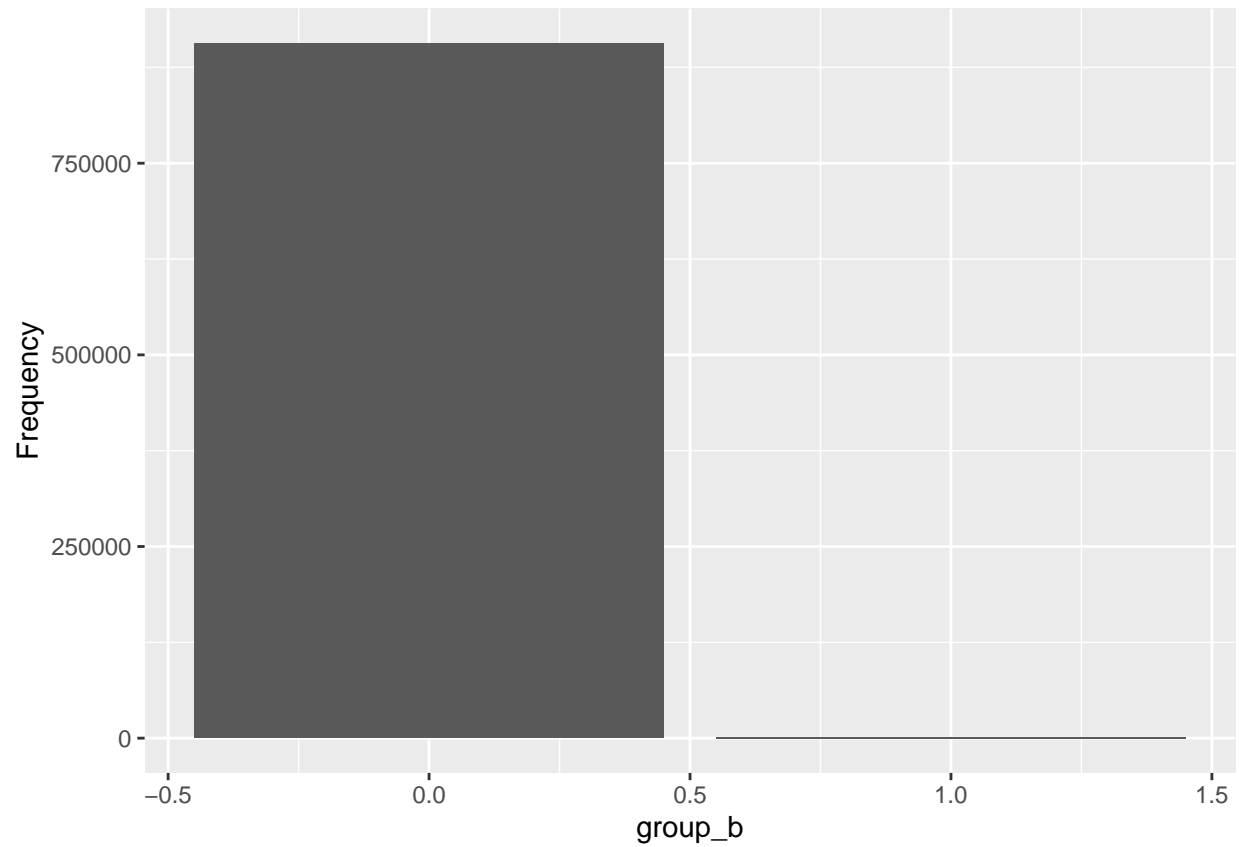
```
[1] should not be used as a predictor
[1] -----
[1] Variable: group_a, type: numeric, requested: yes
[1] Values (2 unique): 0, 1
[1] Missing: 0%
```



```
[1] should not be used as a predictor
[1] -----
[1] Variable: mellon_enr_dt_a, type: character, requested: yes
[1] Values (6 unique): NA, 09/18/2020 00:00:00.000, 09/01/2019 00:00:00.000, 09/01/2022 00:00:00.000, 09/01/2021 00:00:00.000, 09/01/2020 00:00:00.000
[1] Missing: 99.3%
```



```
[1] -----
[1] Variable: group_b, type: numeric, requested: yes
[1] Values (2 unique): 0, 1
[1] Missing: 0%
```



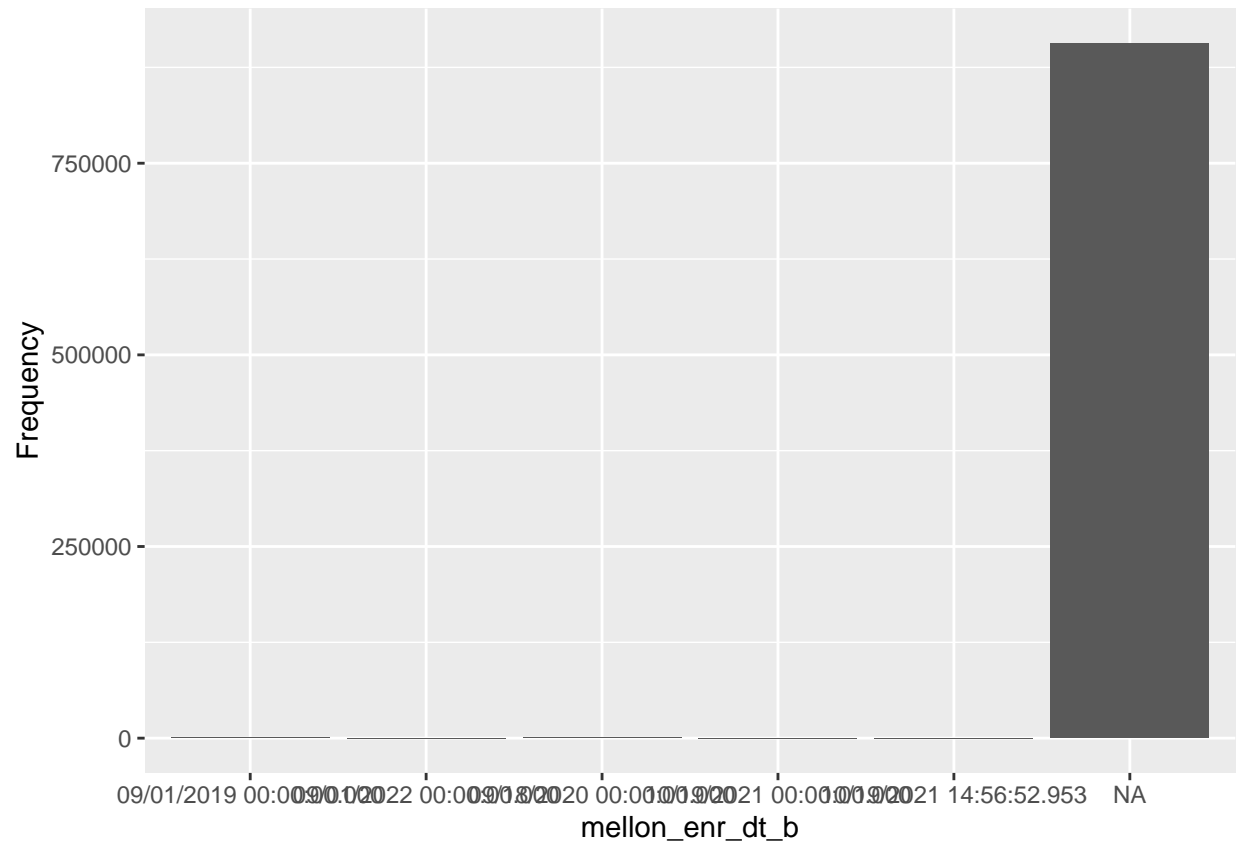
```
[1] should not be used as a predictor
```

```
[1] -----
```

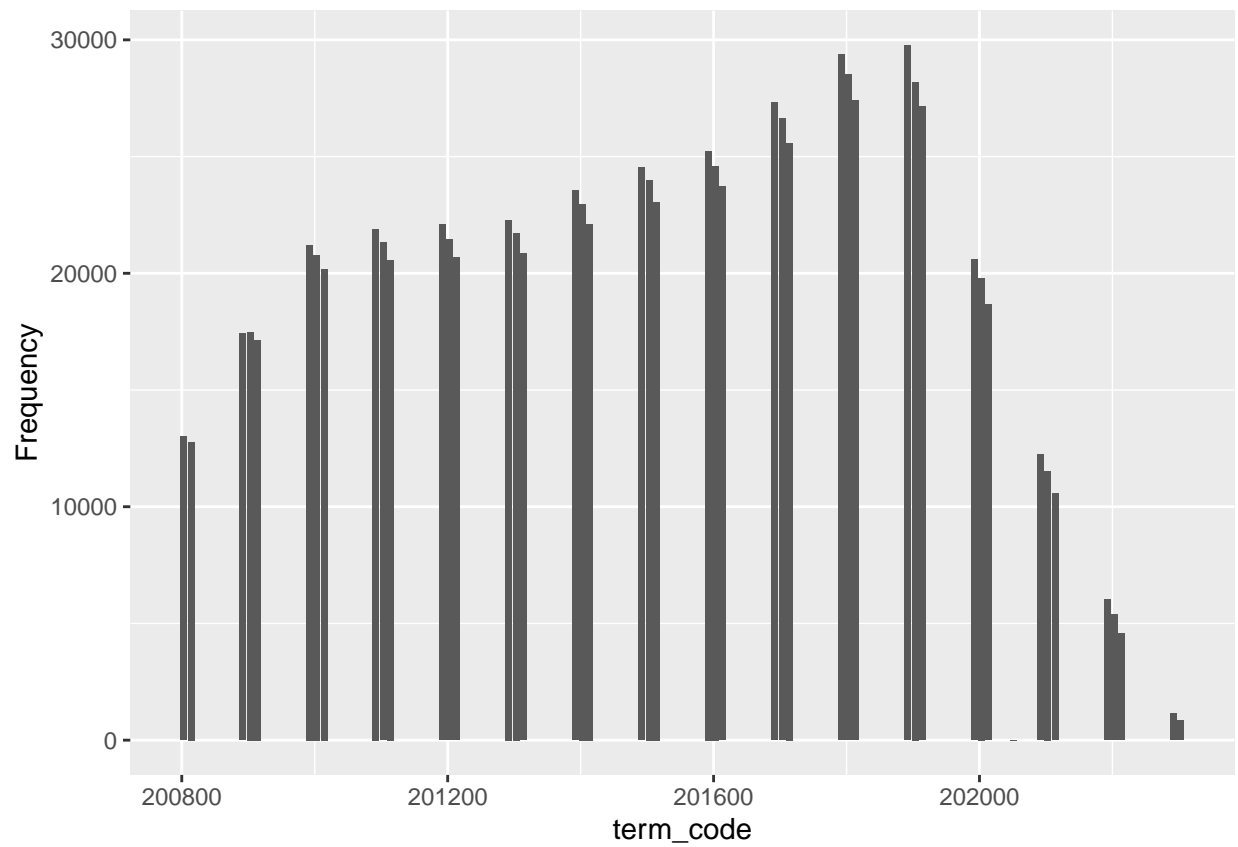
```
[1] Variable: mellon_enr_dt_b, type: character, requested: yes
```

```
[1] Values (6 unique): NA, 09/18/2020 00:00:00.000, 09/01/2019 00:00:00.000, 09/01/2022 00:00:00.000, 1
```

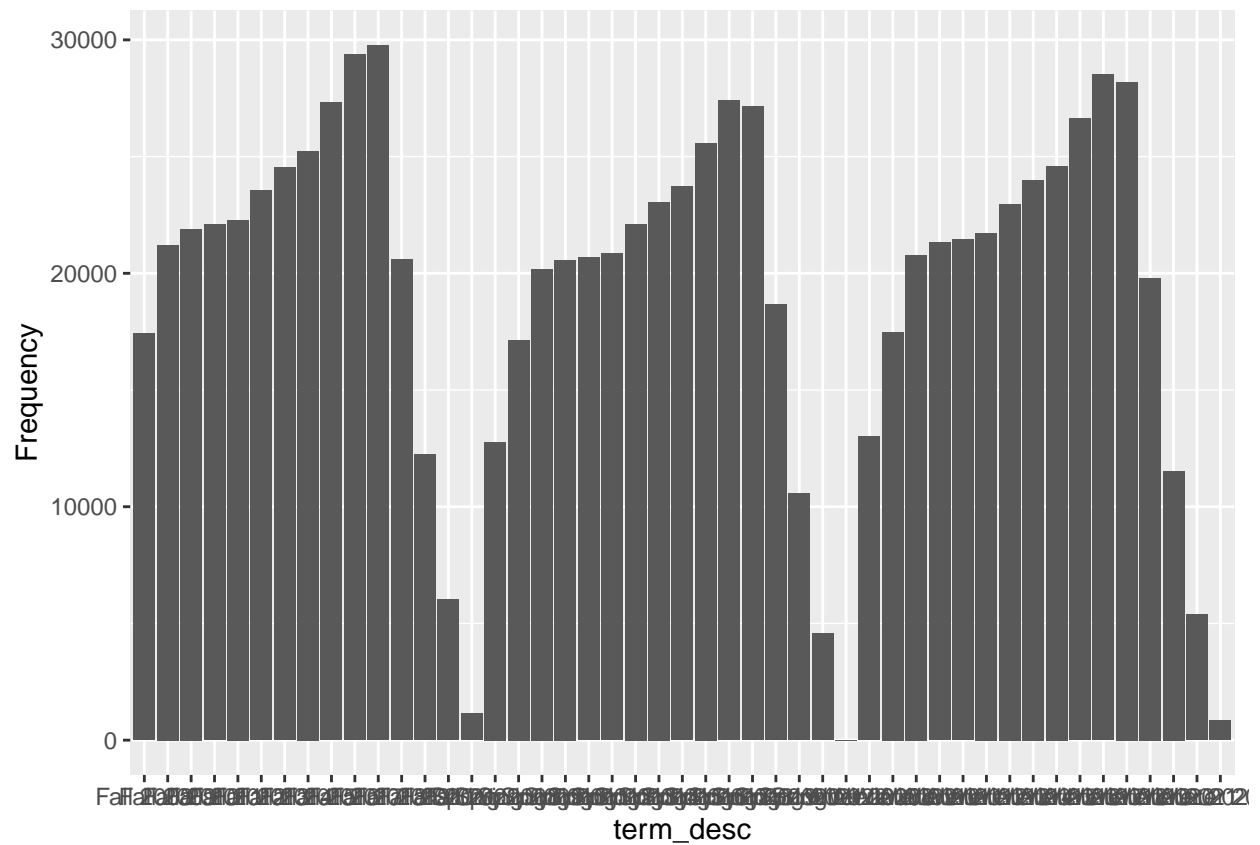
```
[1] Missing: 99.8%
```



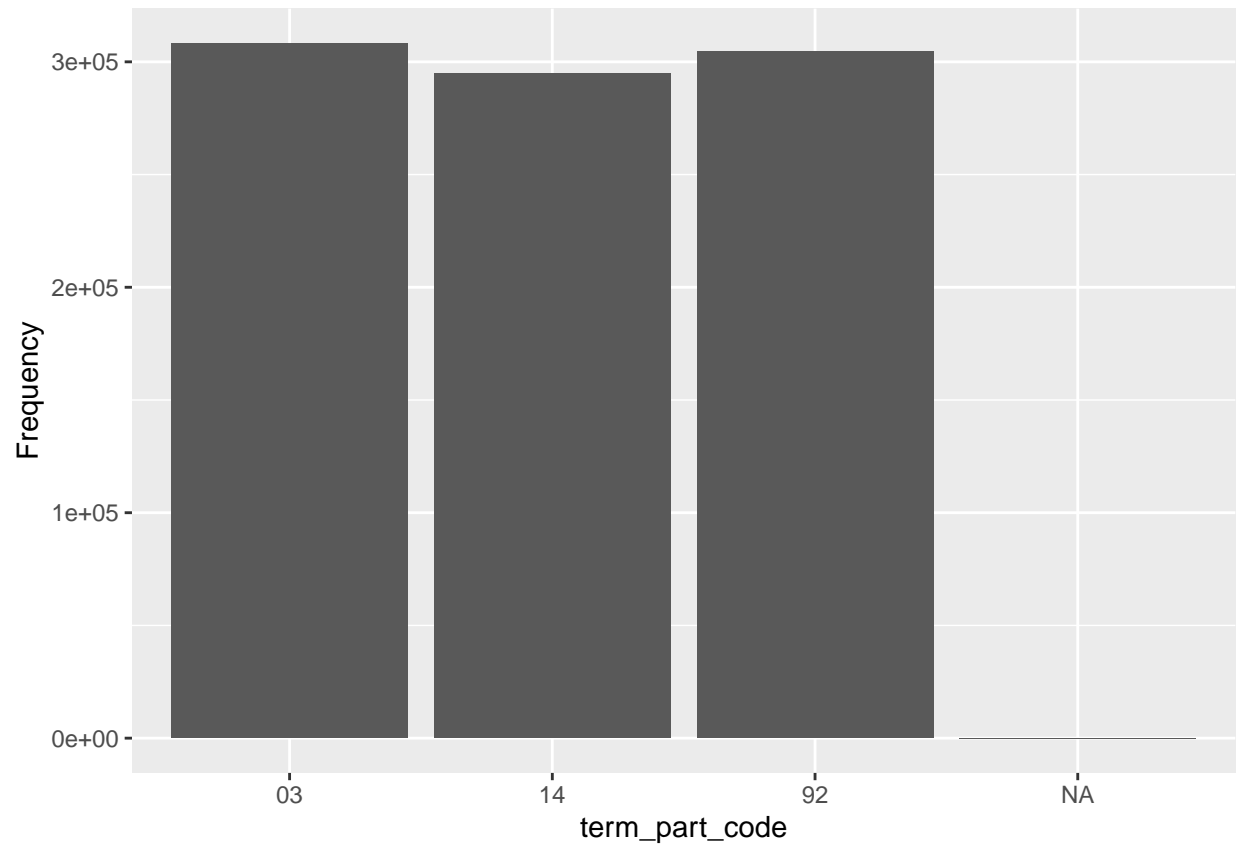
```
[1] -----
[1] Variable: term_code, type: numeric, requested: yes
[1] Values (47 unique): 201892, 202003, 202014, 201914, 202203, ...
[1] Missing: 0%
```



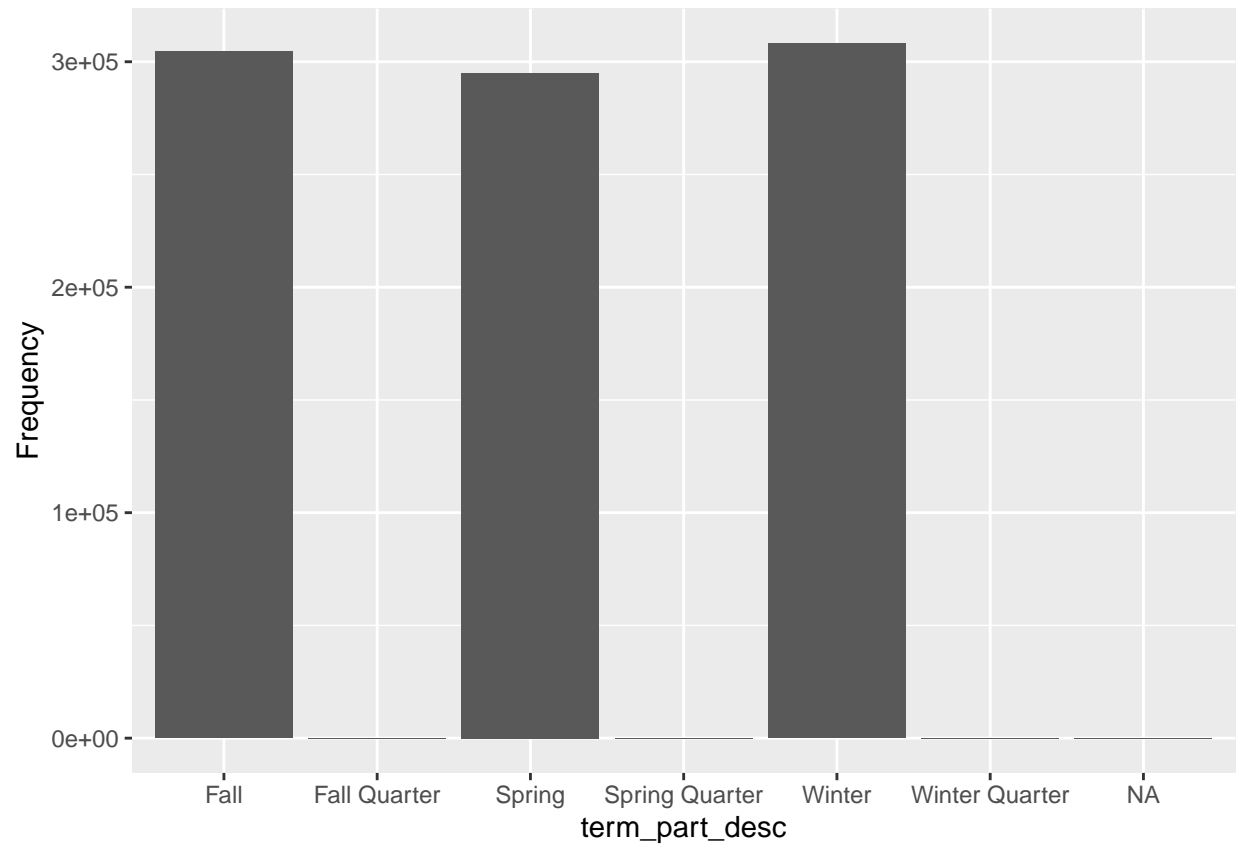
```
[1] -----
[1] Variable: term_desc, type: character, requested: yes
[1] Values (47 unique): Fall 2018, Winter 2020, Spring 2020, Spring 2019, Winter 2022, ...
[1] Missing: 0%
```



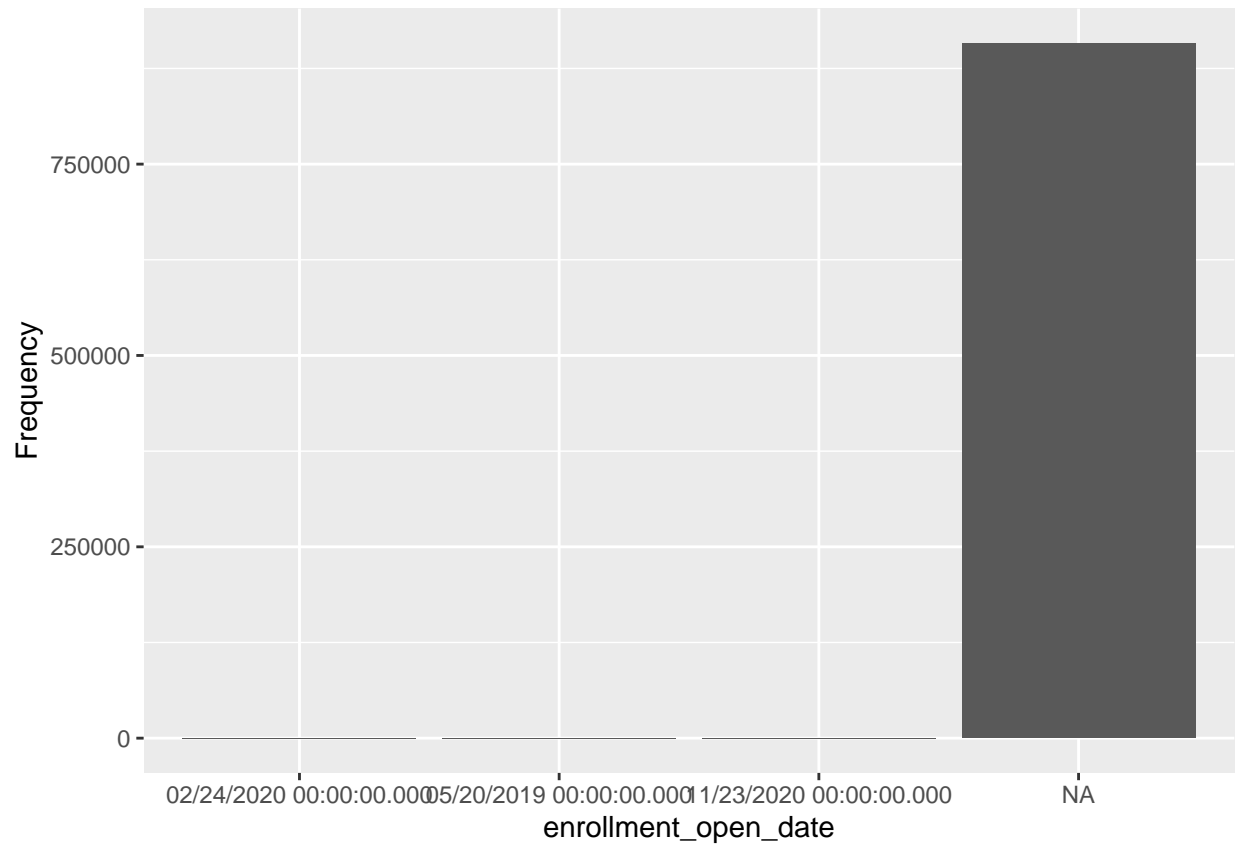
```
[1] -----
[1] Variable: term_part_code, type: character, requested: yes
[1] Values (4 unique): 92, 03, 14, NA
[1] Missing: 0%
```



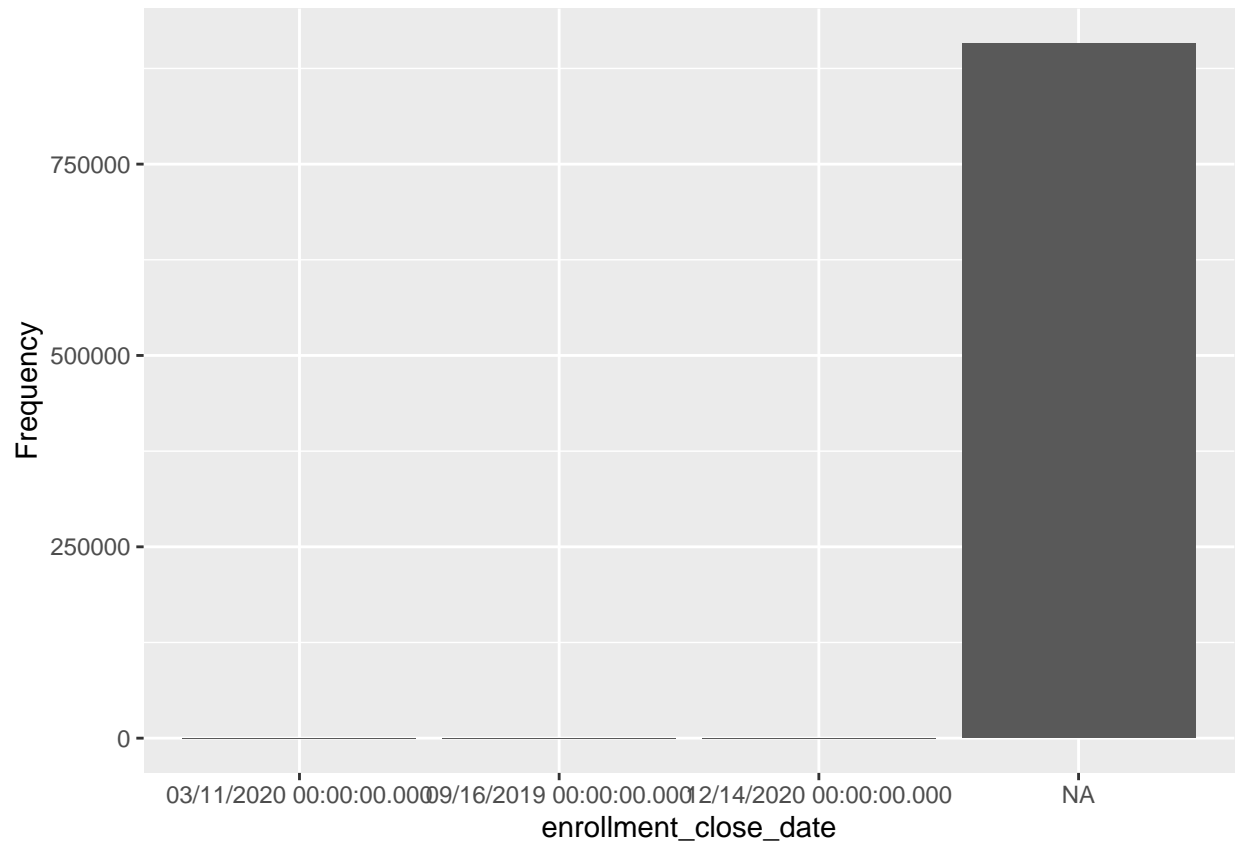
```
[1] -----  
[1] Variable: term_part_desc, type: character, requested: yes  
[1] Values (7 unique): Fall, Winter, Spring, NA, Fall Quarter, ...  
[1] Missing: 0%
```

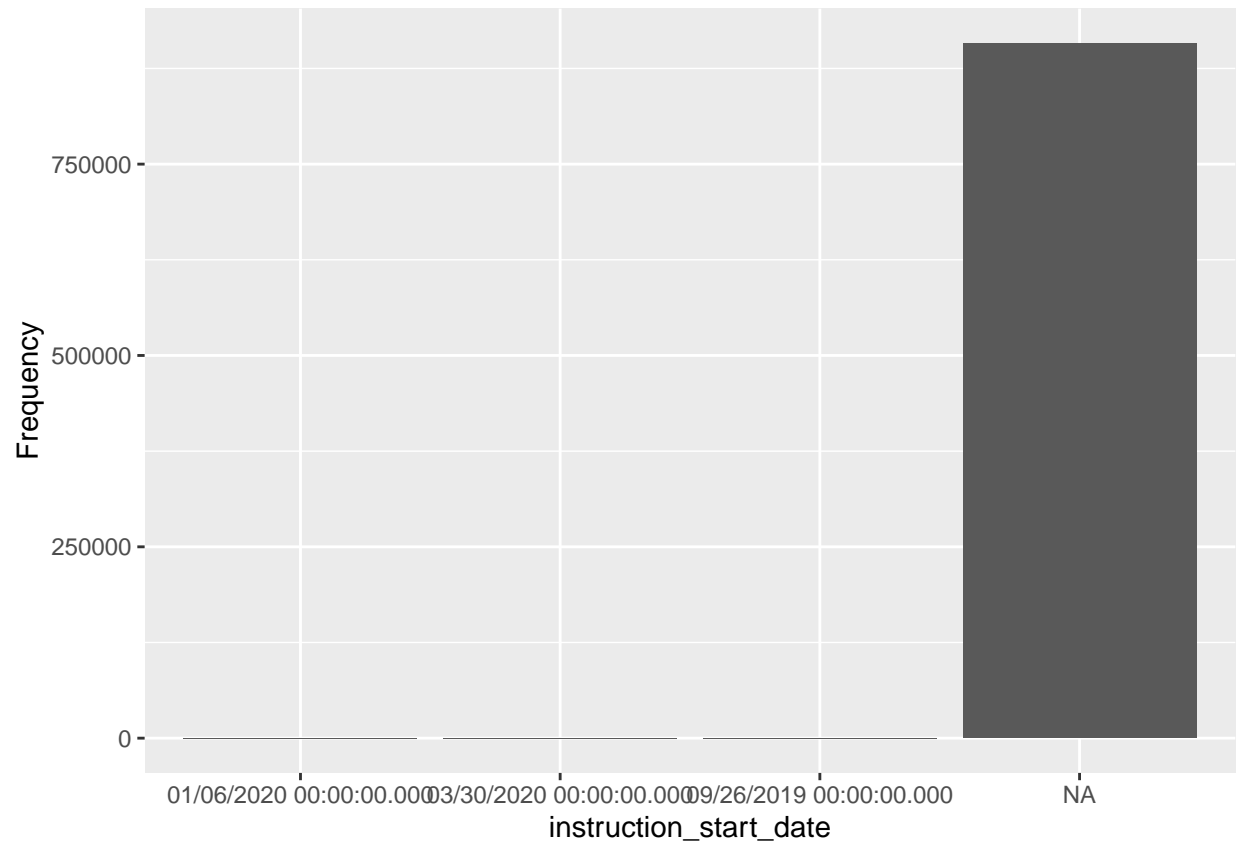
```
[1] -----
[1] Variable: enrollment_open_date, type: character, requested: yes
[1] Values (4 unique): NA, 05/20/2019 00:00:00.000, 11/23/2020 00:00:00.000, 02/24/2020 00:00:00.000
[1] Missing: 100%
```



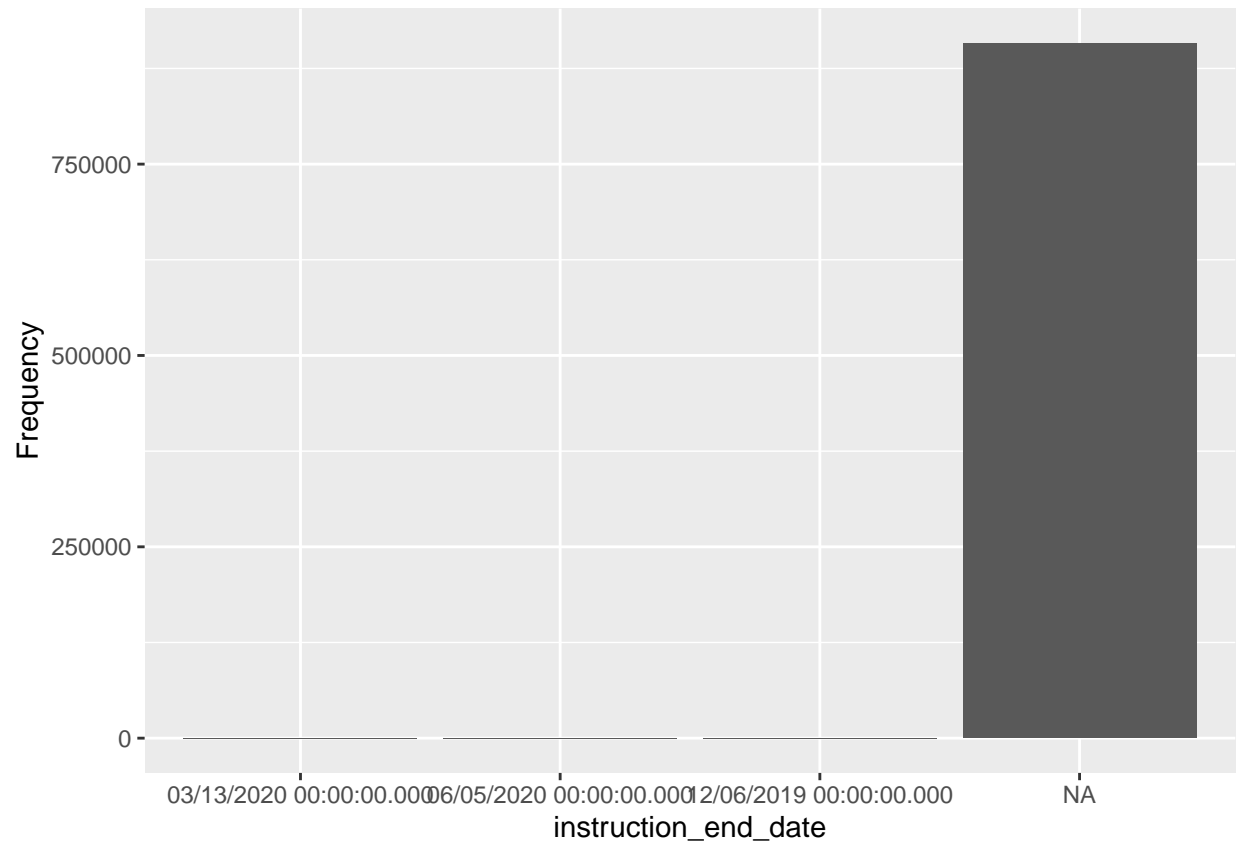
```
[1] -----
[1] Variable: enrollment_close_date, type: character, requested: yes
[1] Values (4 unique): NA, 09/16/2019 00:00:00.000, 12/14/2020 00:00:00.000, 03/11/2020 00:00:00.000
[1] Missing: 100%
```



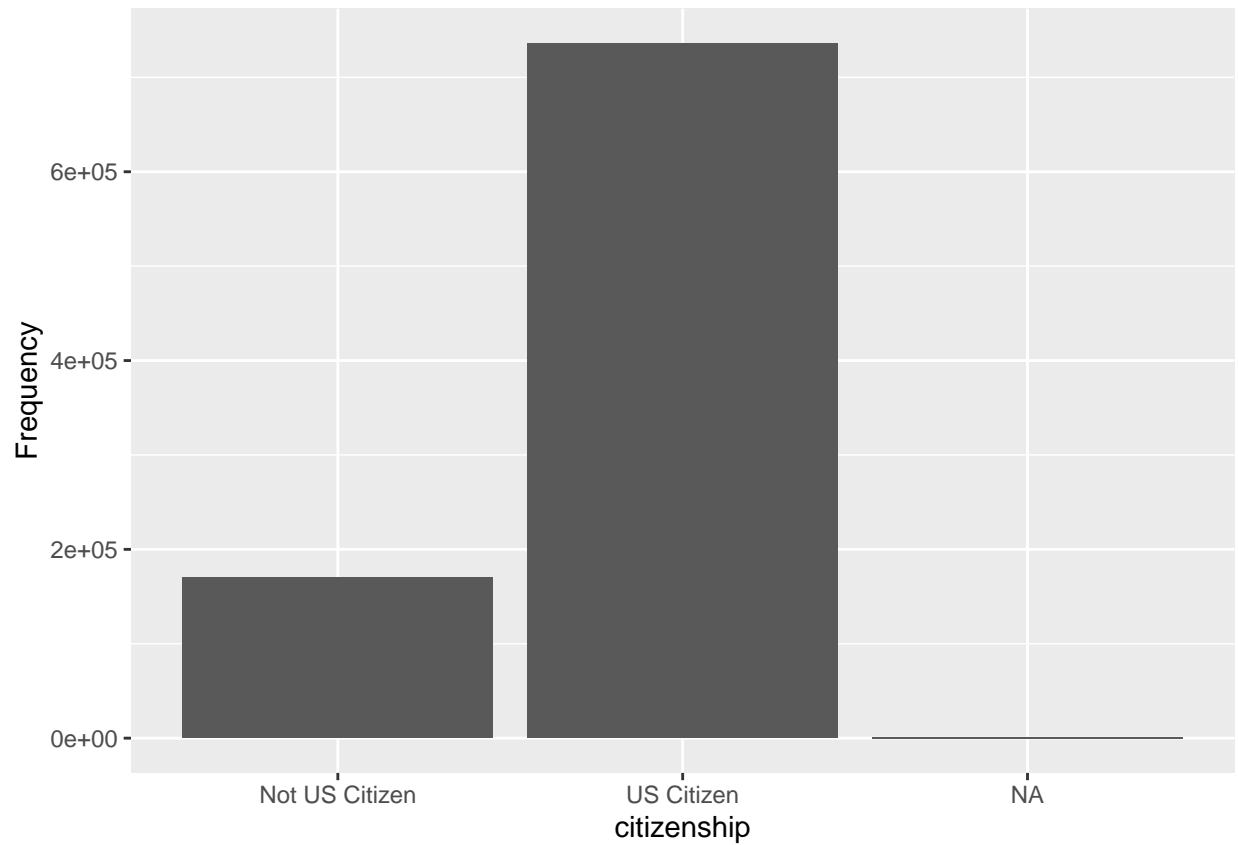
```
[1] -----
[1] Variable: instruction_start_date, type: character, requested: yes
[1] Values (4 unique): NA, 09/26/2019 00:00:00.000, 01/06/2020 00:00:00.000, 03/30/2020 00:00:00.000
[1] Missing: 100%
```



```
[1] -----
[1] Variable: instruction_end_date, type: character, requested: yes
[1] Values (4 unique): NA, 12/06/2019 00:00:00.000, 03/13/2020 00:00:00.000, 06/05/2020 00:00:00.000
[1] Missing: 100%
```

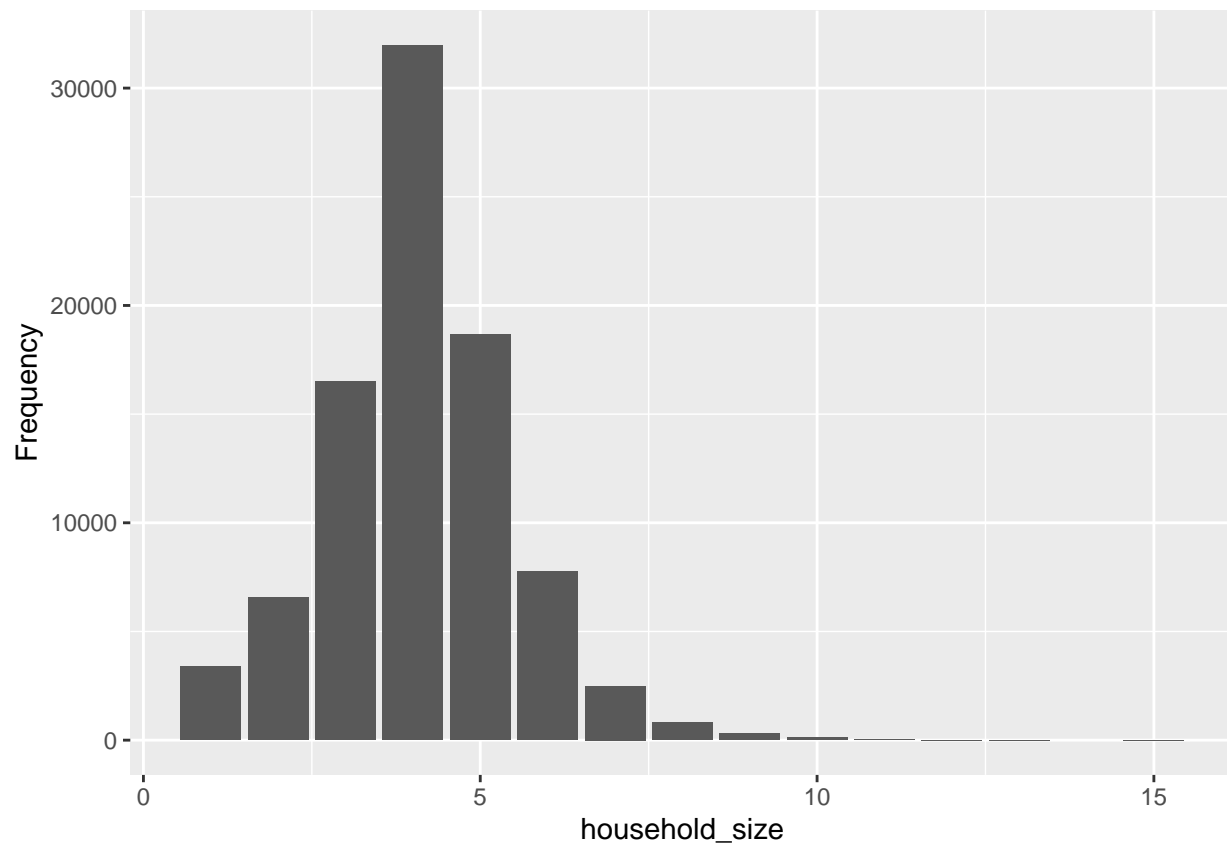


```
[1] -----
[1] Variable: citizenship, type: character, requested: yes
[1] Values (3 unique): US Citizen, Not US Citizen, NA
[1] Missing: 0.1%
```

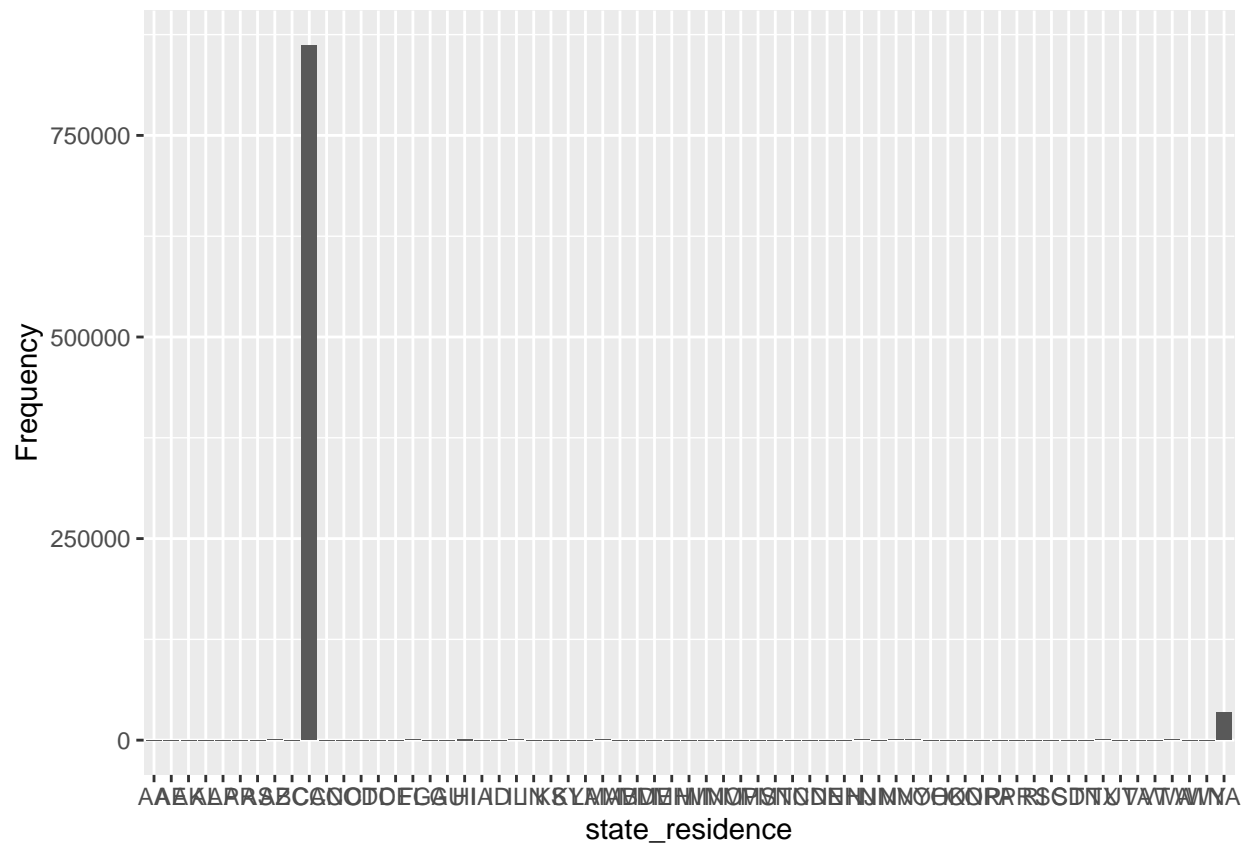


```
[1] -----  
[1] Variable: household_size, type: numeric, requested: yes  
[1] Values (15 unique): NA, 3, 4, 6, 5, ...  
[1] Missing: 90.2%  
[1] Most missing: 2018 100%, Least missing: 2015 67.1%
```

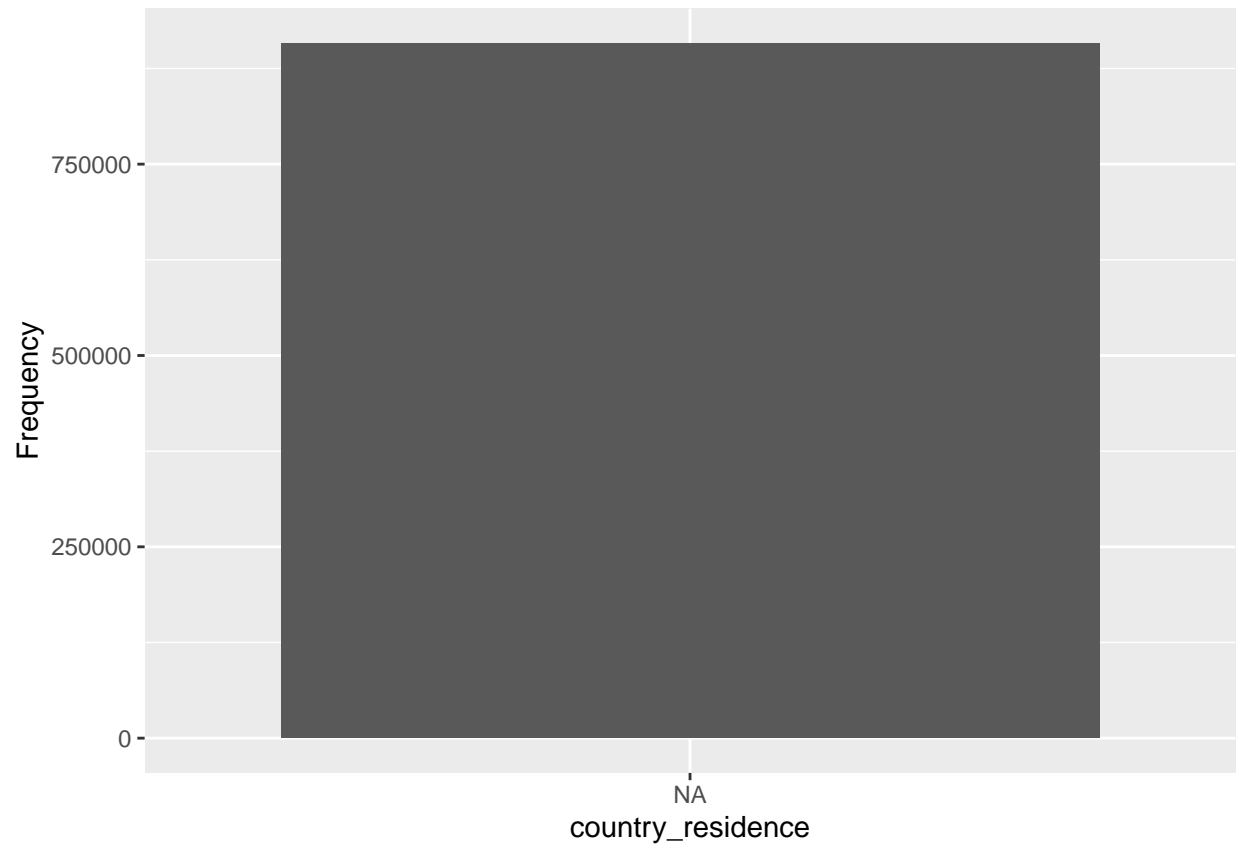
Warning: Removed 819234 rows containing non-finite values ('stat_count()').



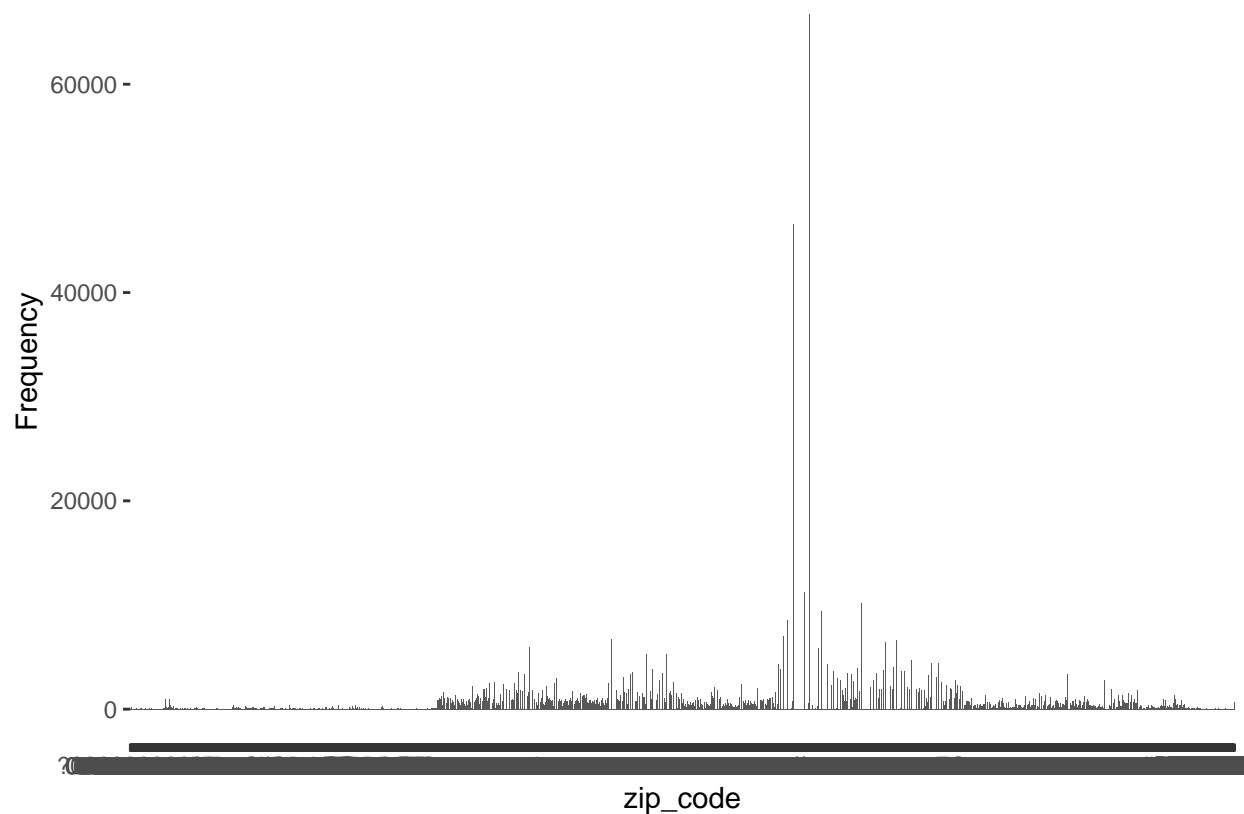
```
[1] -----
[1] Variable: city_residence, type: character, requested: yes
[1] Values (3464 unique): HUNTINGTON BEACH, OCEANSIDE, SANTA ANA, SAN DIMAS, WEST COVINA, ...
[1] Missing: 0%
[1] -----
[1] Variable: state_residence, type: character, requested: yes
[1] Values (63 unique): CA, NV, NA, RI, CT, ...
[1] Missing: 3.8%
[1] Most missing: 2016 9.6%, Least missing: 2006 0.1%
```



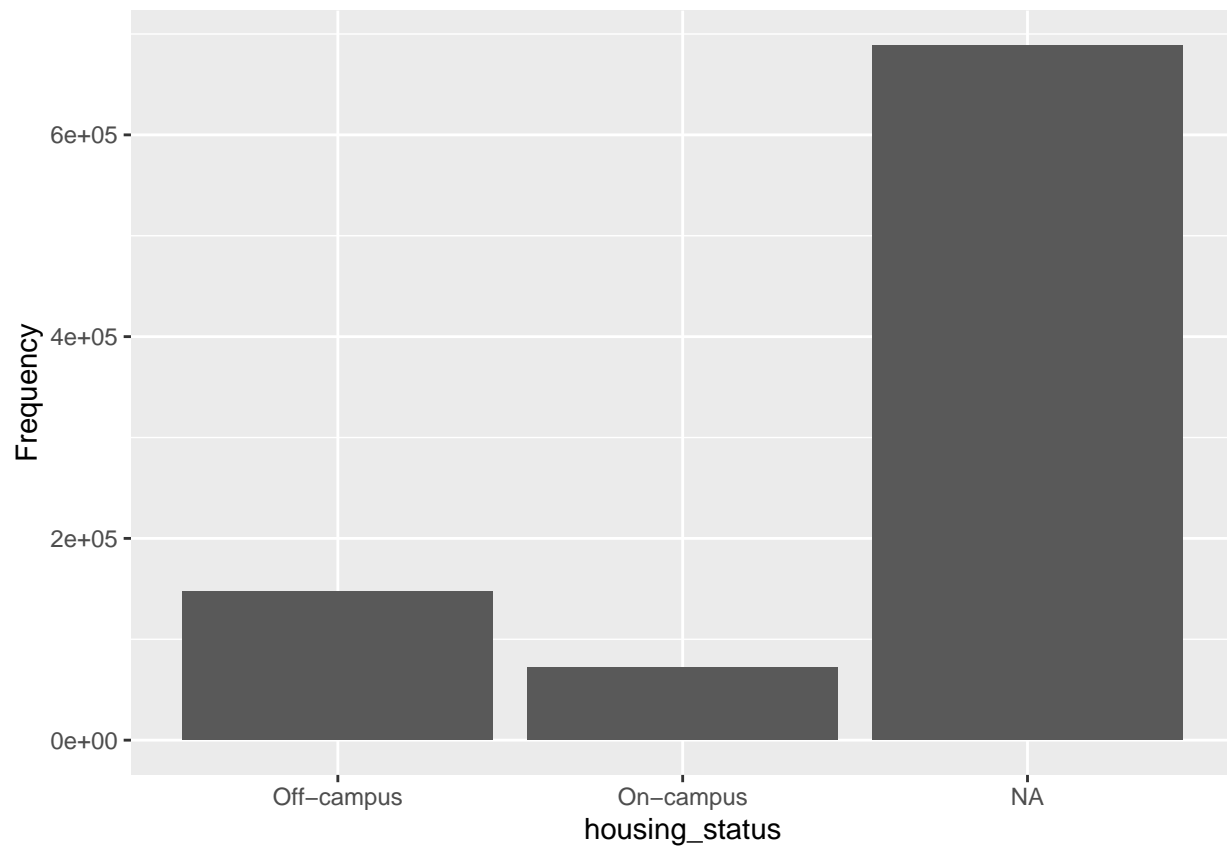
```
[1] -----
[1] Variable: country_residence, type: logical, requested: yes
[1] Values (1 unique): NA
[1] Missing: 100%
```

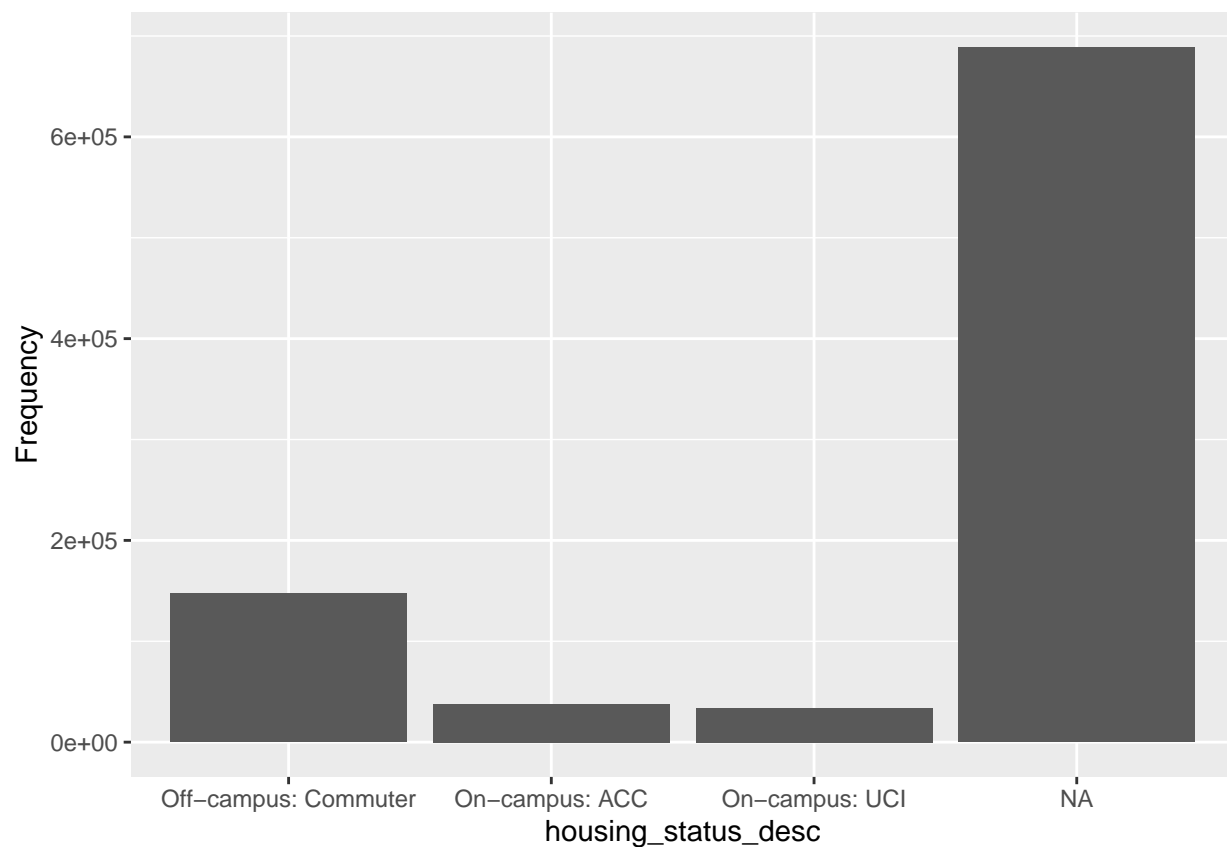
```
[1] -----  
[1] Variable: zip_code, type: character, requested: yes  
[1] Values (10975 unique): 92648, 92057, 92707, 91773, 91791, ...  
[1] Missing: 0.1%
```



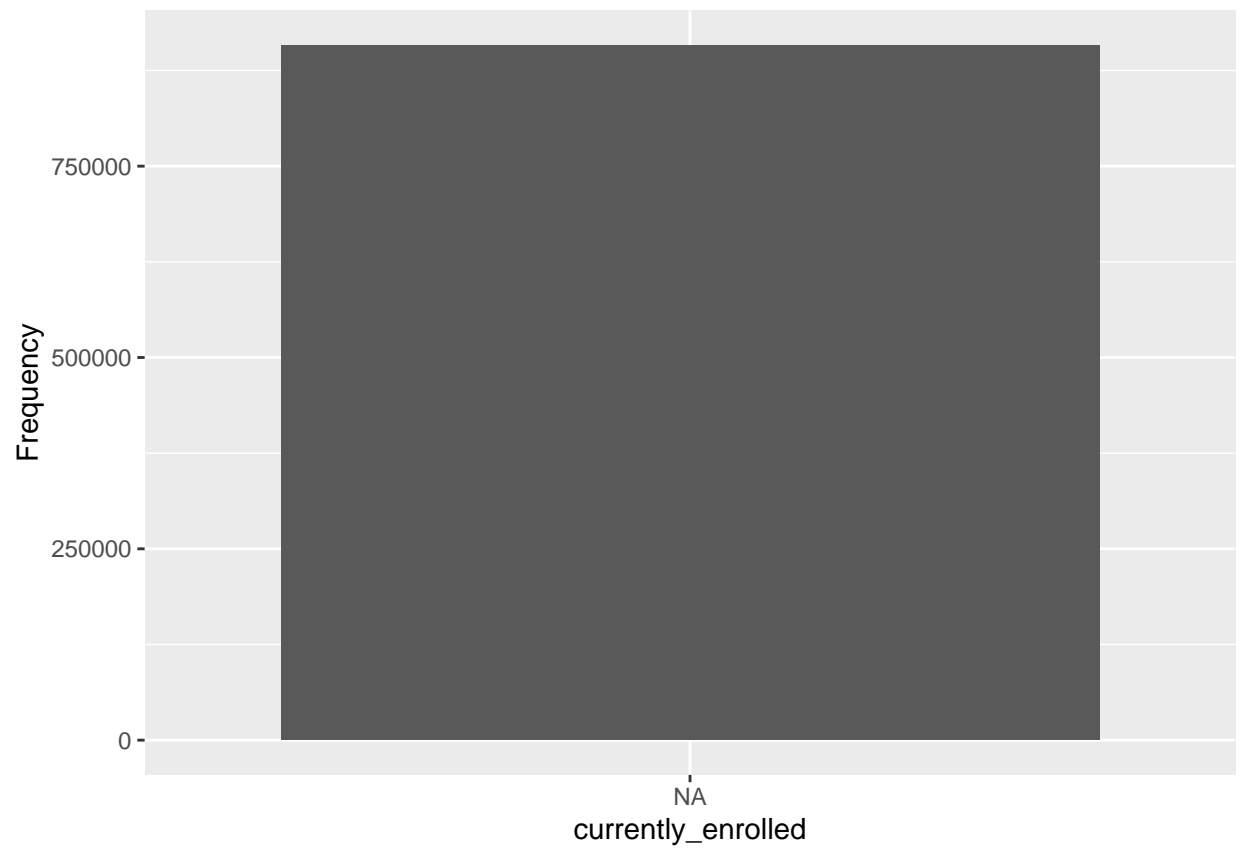
```
[1] -----
[1] Variable: housing_status, type: character, requested: yes
[1] Values (3 unique): NA, On-campus, Off-campus
[1] Missing: 75.9%
  Group.1      x
1    2006 0.9989400
2    2007 0.9995472
3    2008 0.9993435
4    2009 0.9988065
5    2010 0.9954115
6    2011 0.9655814
7    2012 0.8860318
8    2013 0.8267333
9    2014 0.7459012
10   2015 0.6132671
11   2016 0.6784658
12   2017 0.4386285
13   2018 0.1405510
```



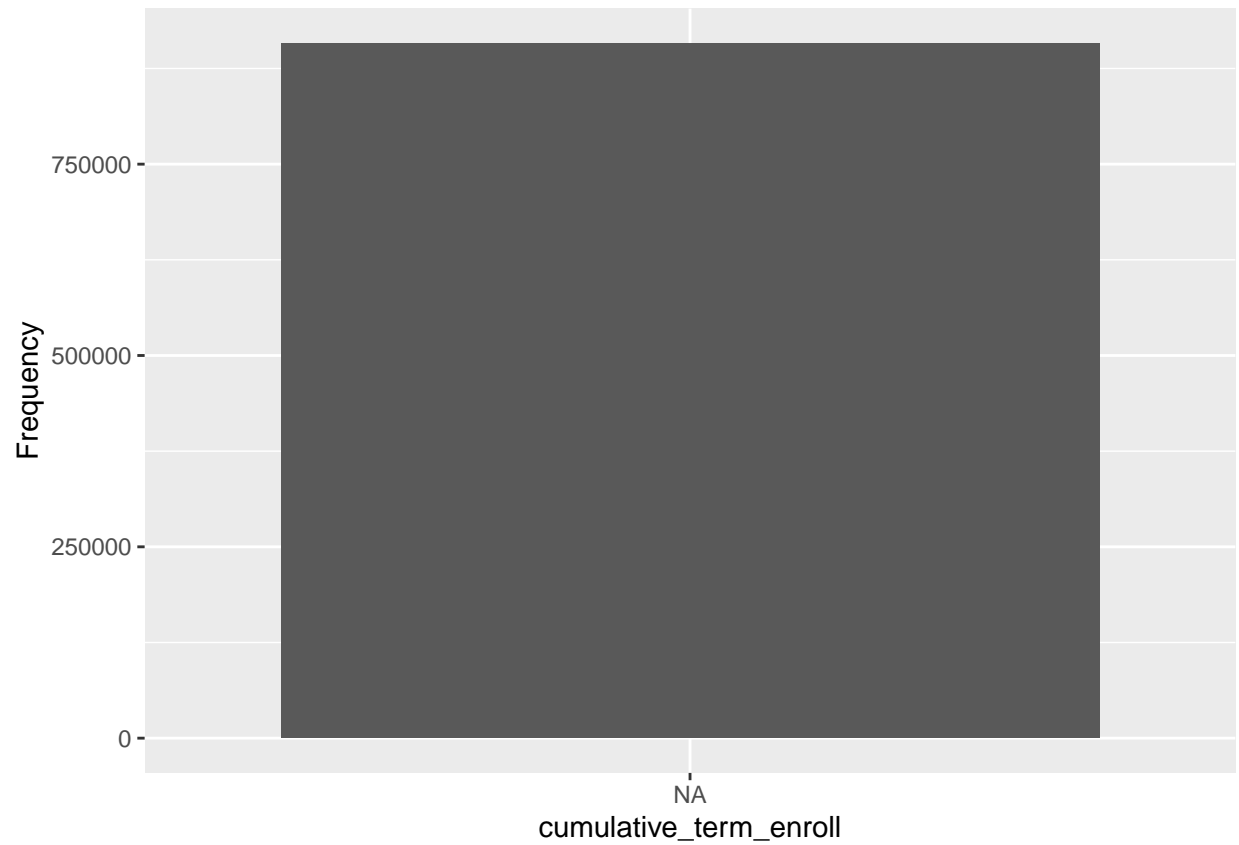
```
[1] -----
[1] Variable: housing_status_desc, type: character, requested: yes
[1] Values (4 unique): NA, On-campus: ACC, Off-campus: Commuter, On-campus: UCI
[1] Missing: 75.9%
  Group.1      x
1    2006 0.9989400
2    2007 0.9995472
3    2008 0.9993435
4    2009 0.9988065
5    2010 0.9954115
6    2011 0.9655814
7    2012 0.8860318
8    2013 0.8267333
9    2014 0.7459012
10   2015 0.6132671
11   2016 0.6784658
12   2017 0.4386285
13   2018 0.1405510
```



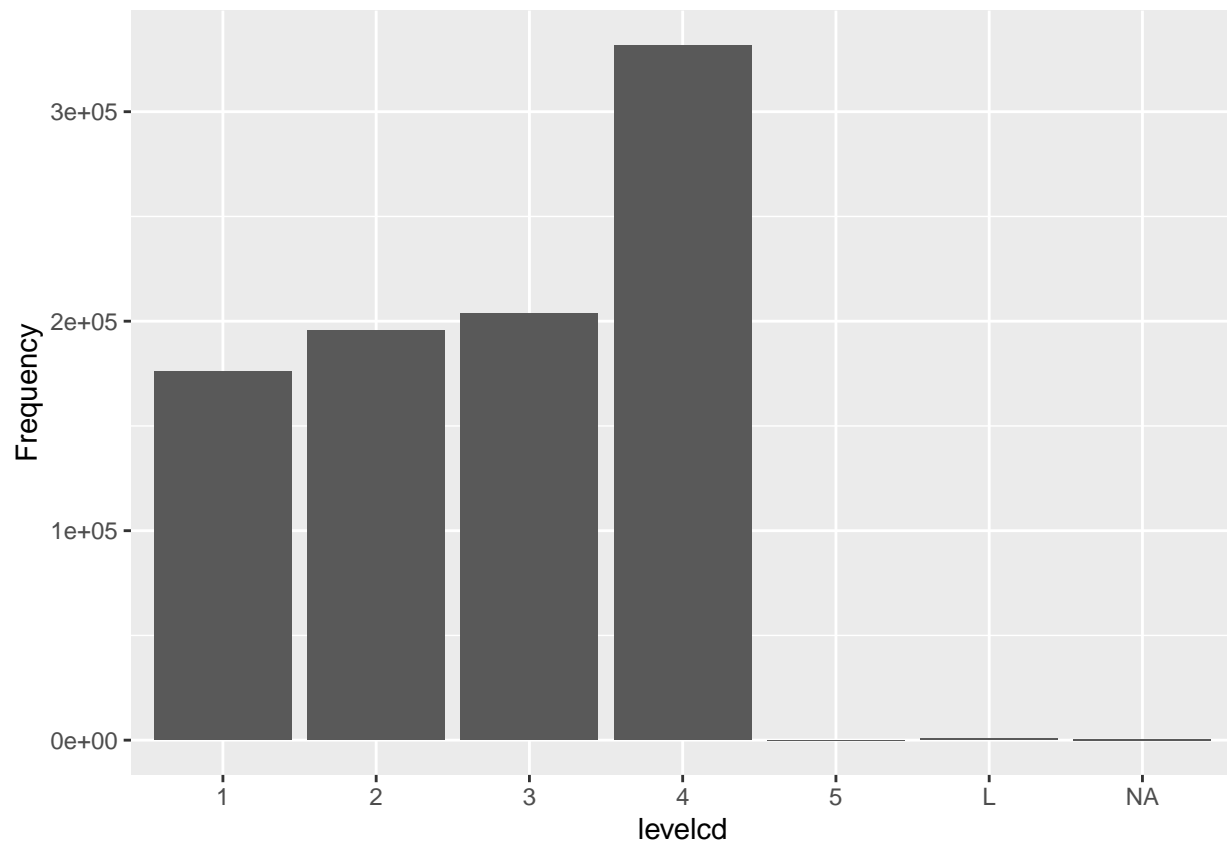
```
[1] -----
[1] Variable: housing_complex_name, type: character, requested: yes
[1] Values (15 unique): NA, Vista del Campo, Off-campus commuter, Middle Earth, Mesa Court, ...
[1] Missing: 75.9%
  Group.1      x
1    2006 0.9989400
2    2007 0.9995472
3    2008 0.9993435
4    2009 0.9988065
5    2010 0.9954115
6    2011 0.9655814
7    2012 0.8860318
8    2013 0.8267333
9    2014 0.7459012
10   2015 0.6132671
11   2016 0.6784658
12   2017 0.4386285
13   2018 0.1405510
```

```
[1] -----  
[1] Variable: cumulative_term_enroll, type: logical, requested: yes  
[1] Values (1 unique): NA  
[1] Missing: 100%
```

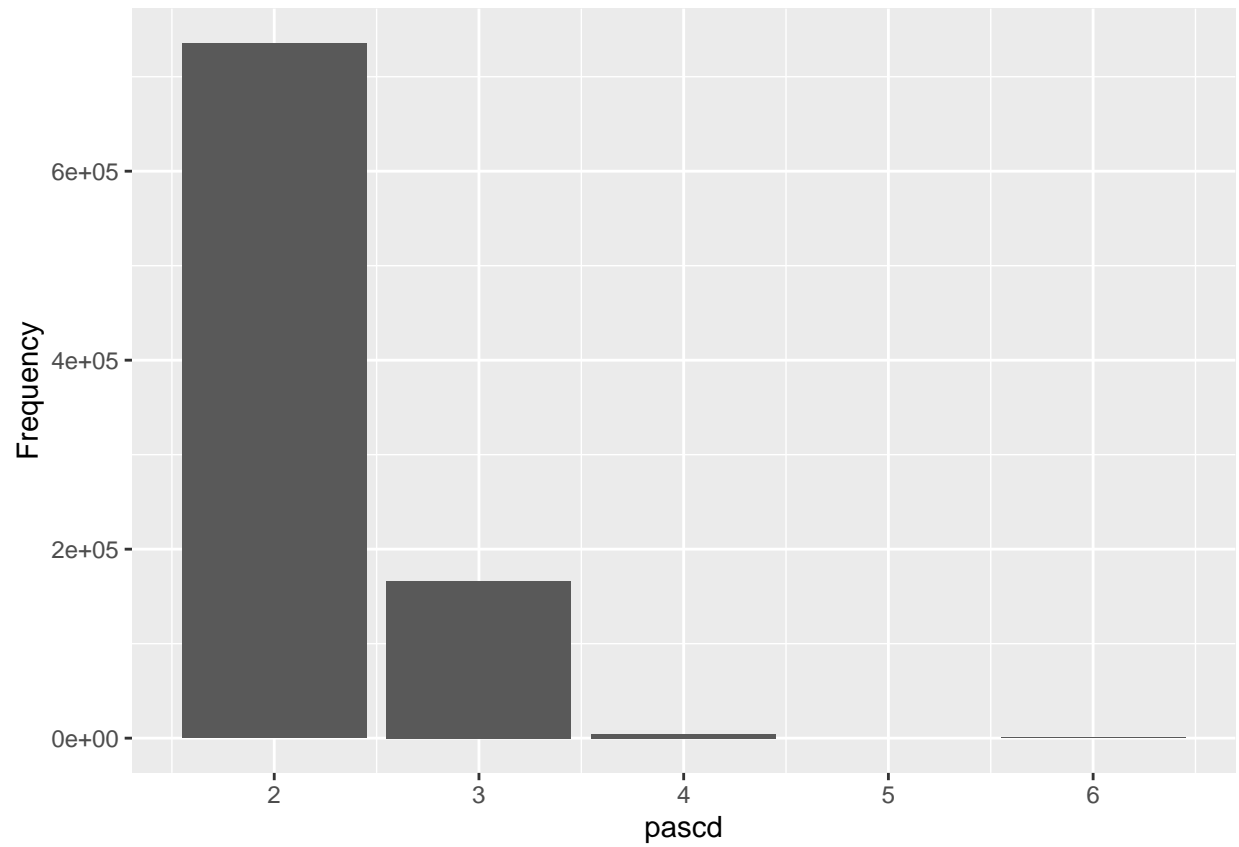


```
[1] -----  
[1] Variable: levelcd, type: character, requested: yes  
[1] Values (7 unique): 1, 3, 4, 2, NA, ...  
[1] Missing: 0.1%
```

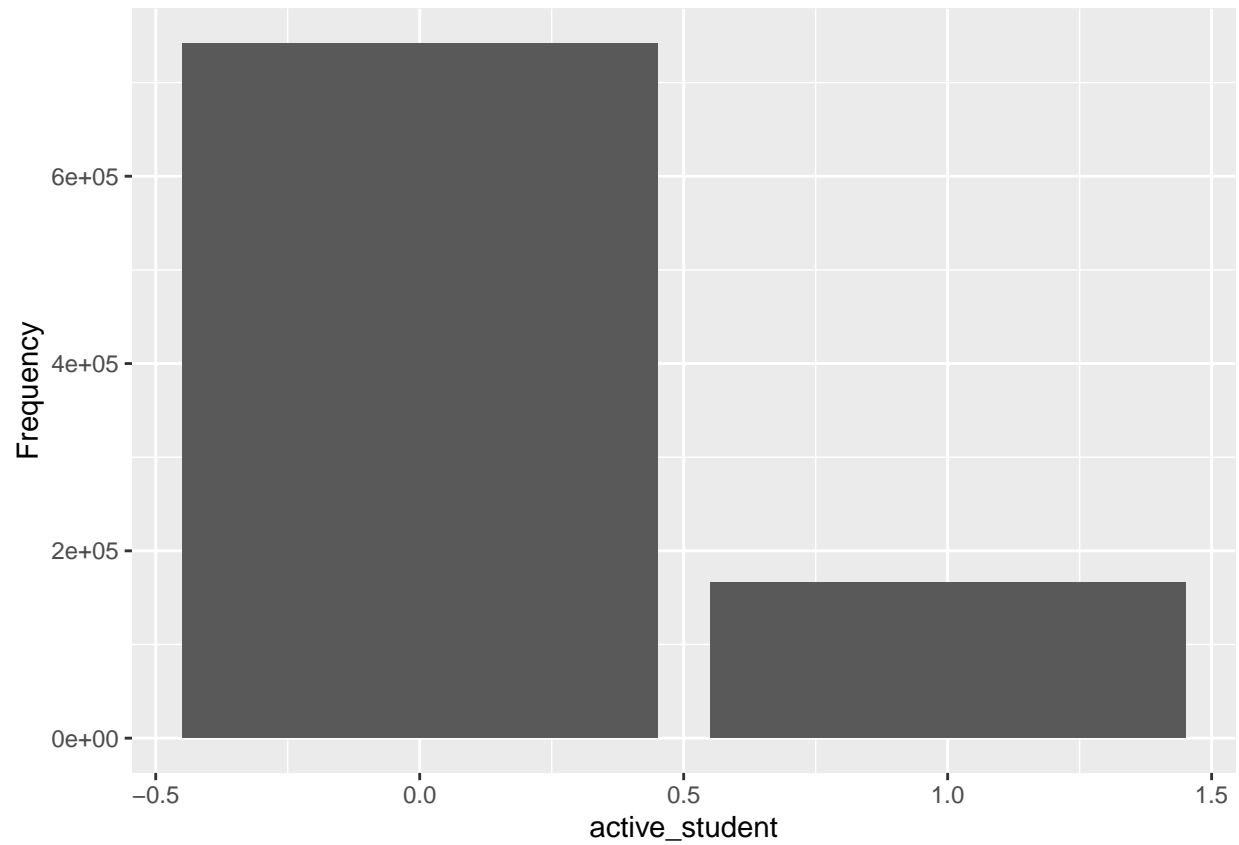


```
[1] -----  
[1] Variable: pascd, type: numeric, requested: yes  
[1] Values (5 unique): 2, 3, 4, NA, 6  
[1] Missing: 0.1%
```

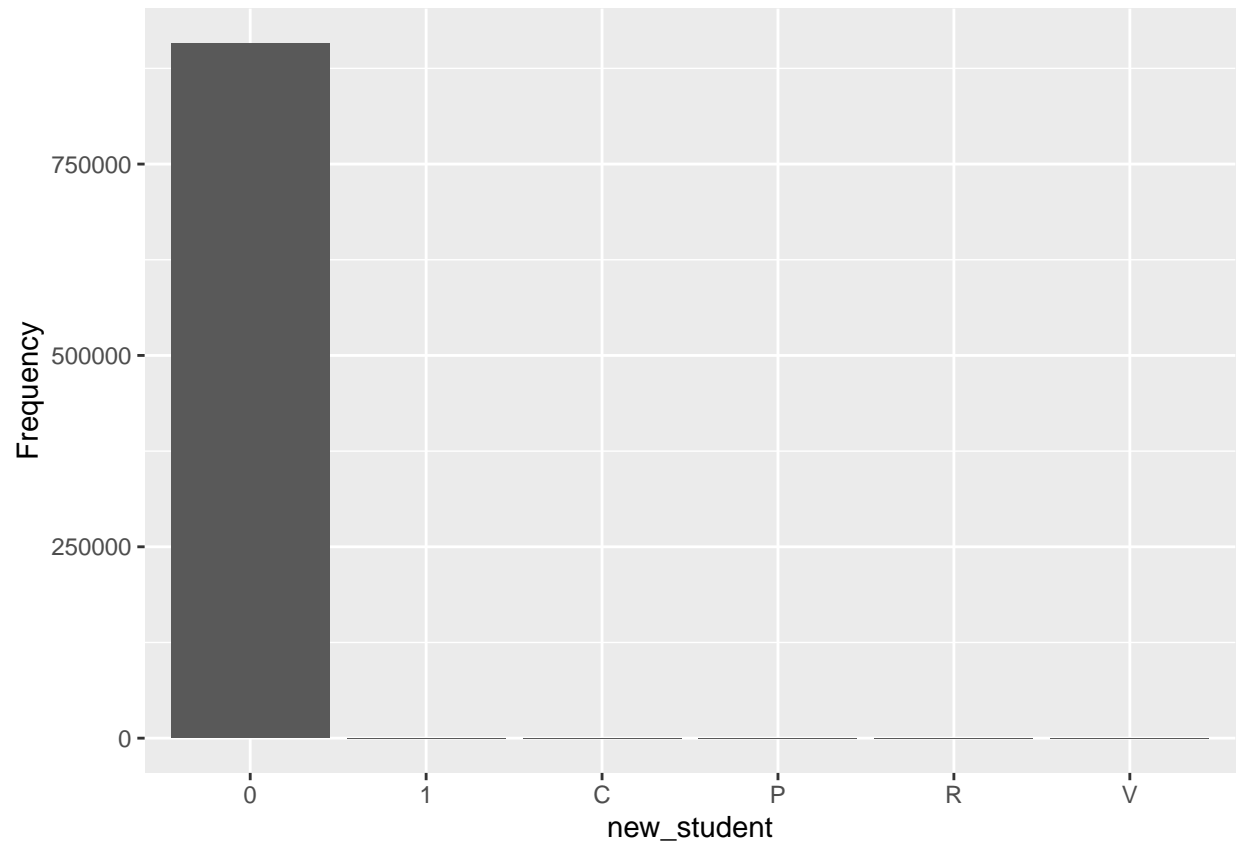
Warning: Removed 964 rows containing non-finite values ('stat_count()').



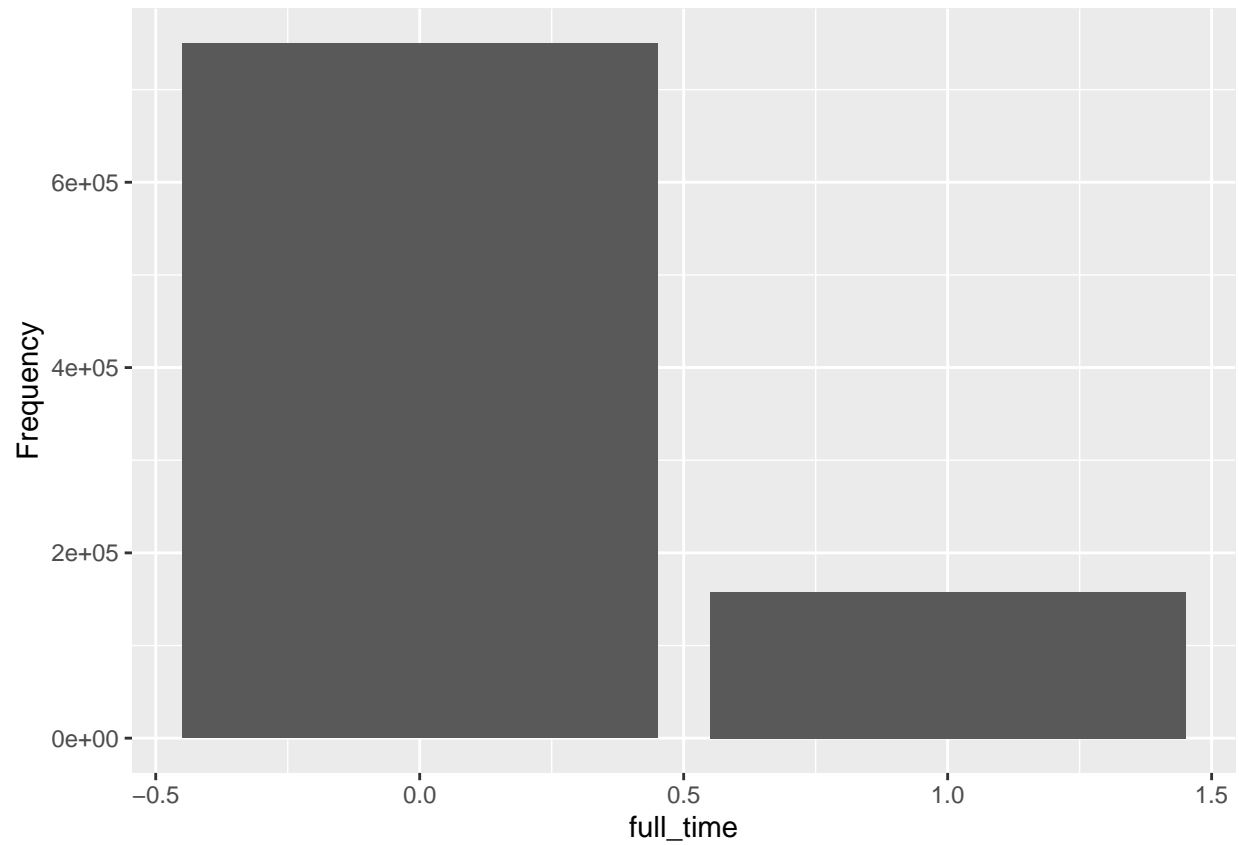
```
[1] -----  
[1] Variable: active_student, type: numeric, requested: yes  
[1] Values (2 unique): 0, 1  
[1] Missing: 0%
```



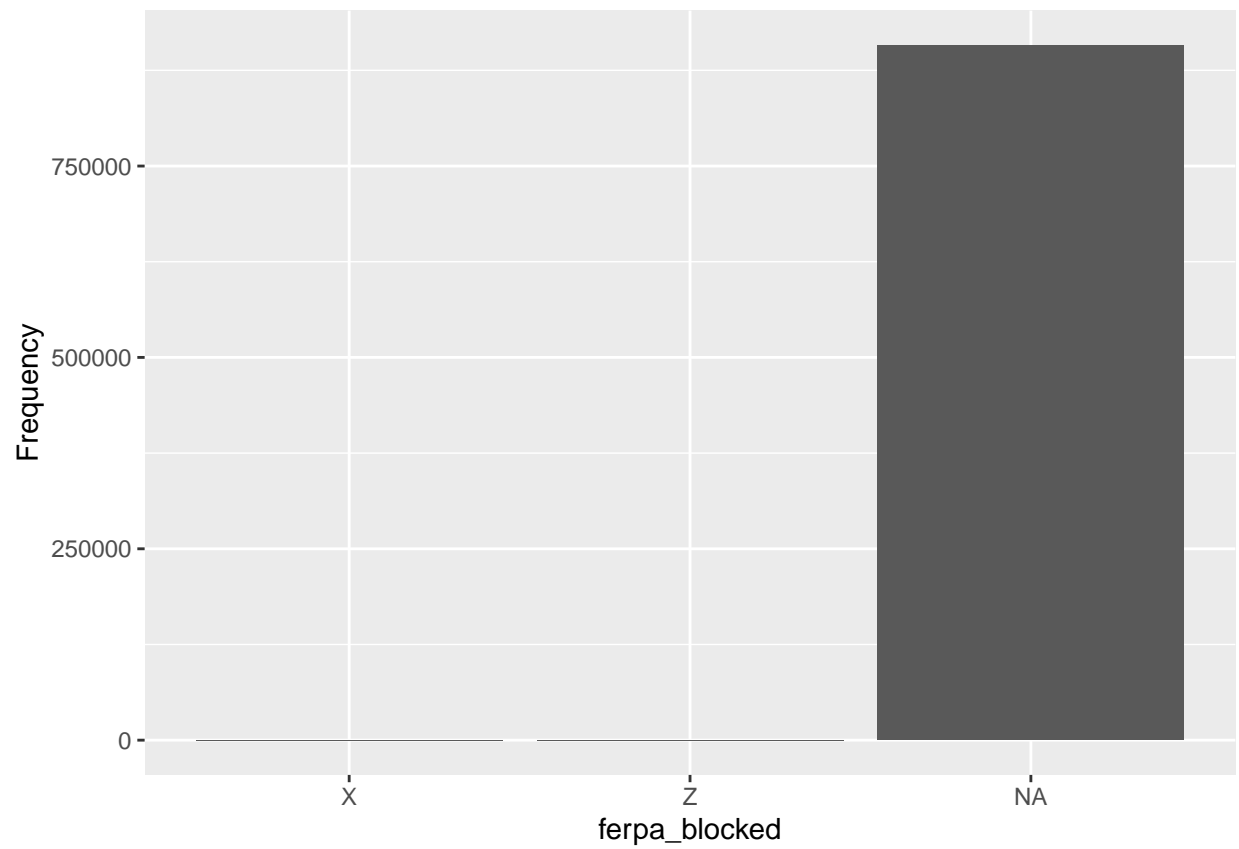
```
[1] -----  
[1] Variable: new_student, type: character, requested: yes  
[1] Values (6 unique): 0, 1, C, R, P, ...  
[1] Missing: 0%
```



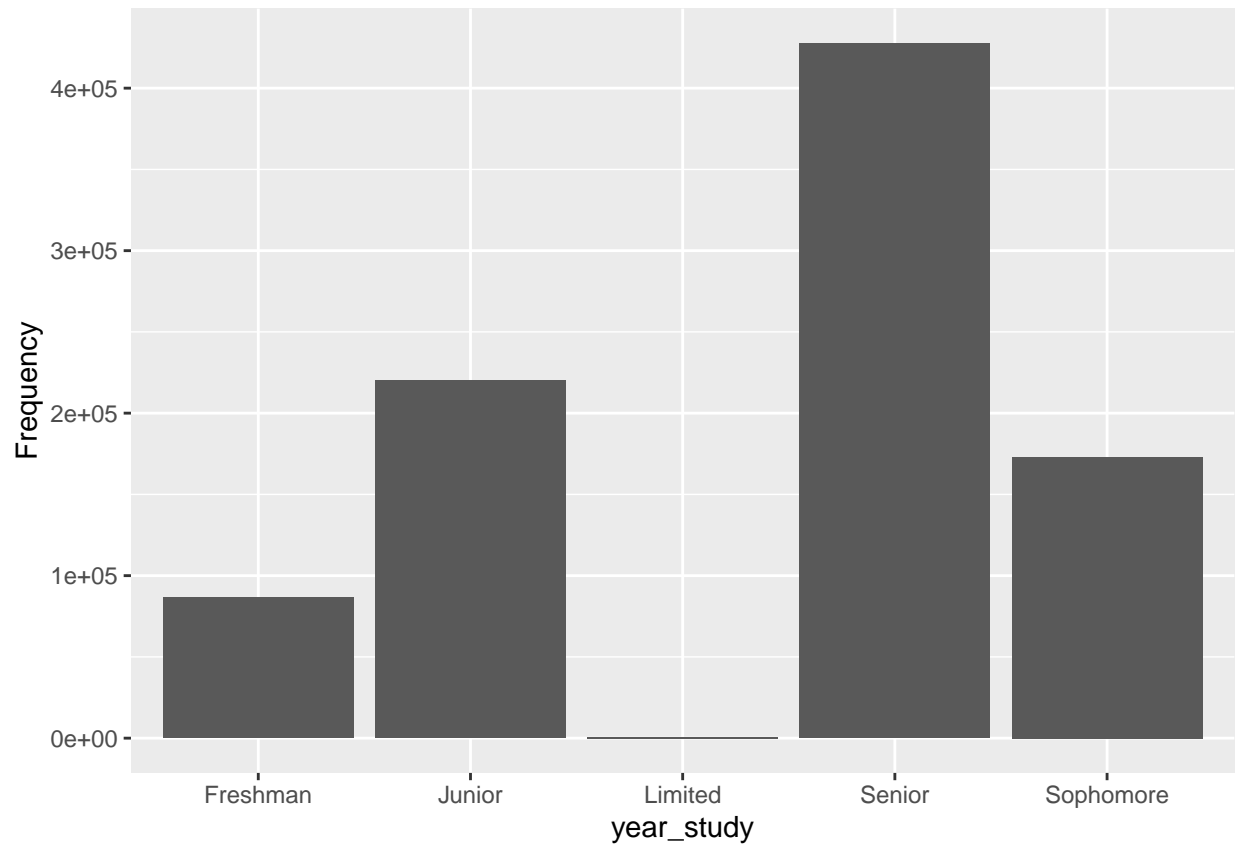
```
[1] -----  
[1] Variable: full_time, type: numeric, requested: yes  
[1] Values (2 unique): 0, 1  
[1] Missing: 0%
```



```
[1] -----  
[1] Variable: ferpa_blocked, type: character, requested: yes  
[1] Values (3 unique): NA, X, Z  
[1] Missing: 100%
```

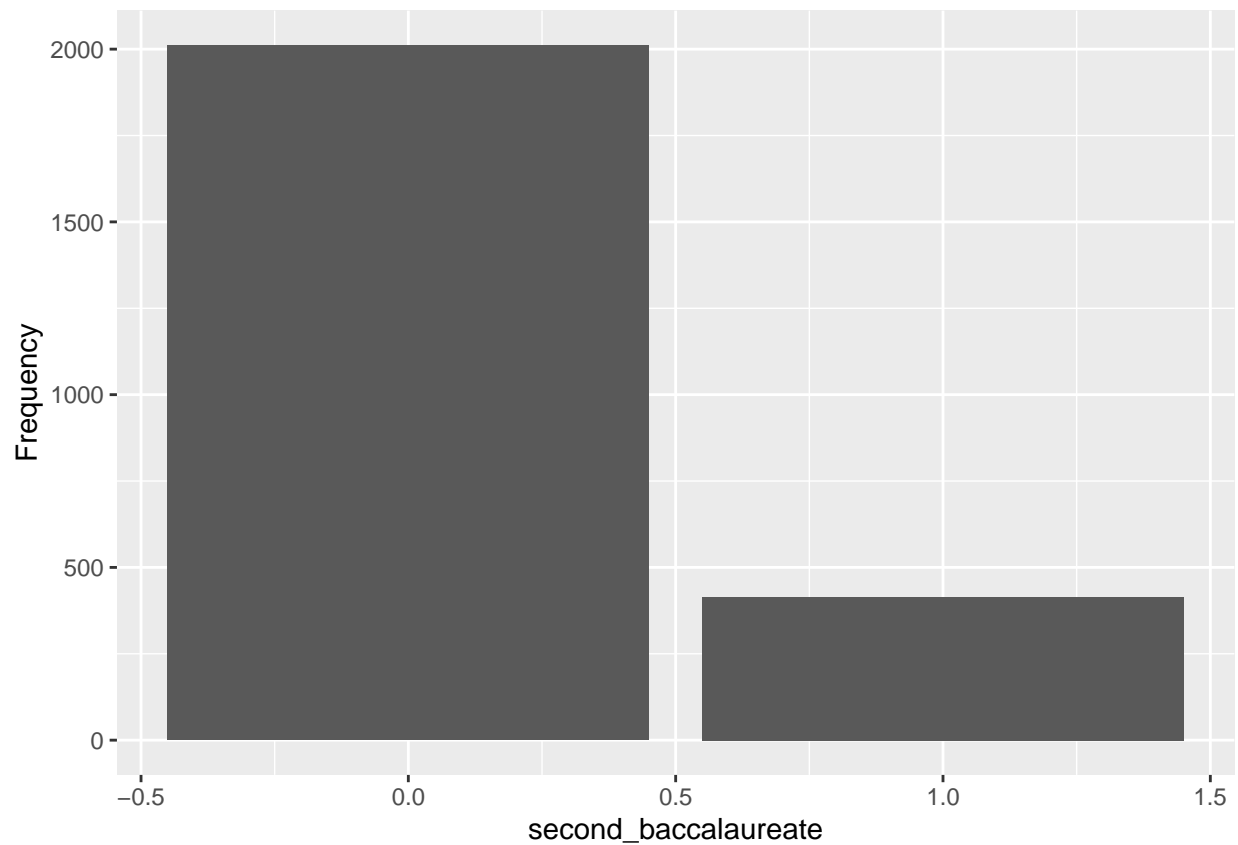


```
[1] -----  
[1] Variable: year_study, type: character, requested: yes  
[1] Values (5 unique): Freshman, Sophomore, Junior, Senior, Limited  
[1] Missing: 0%
```



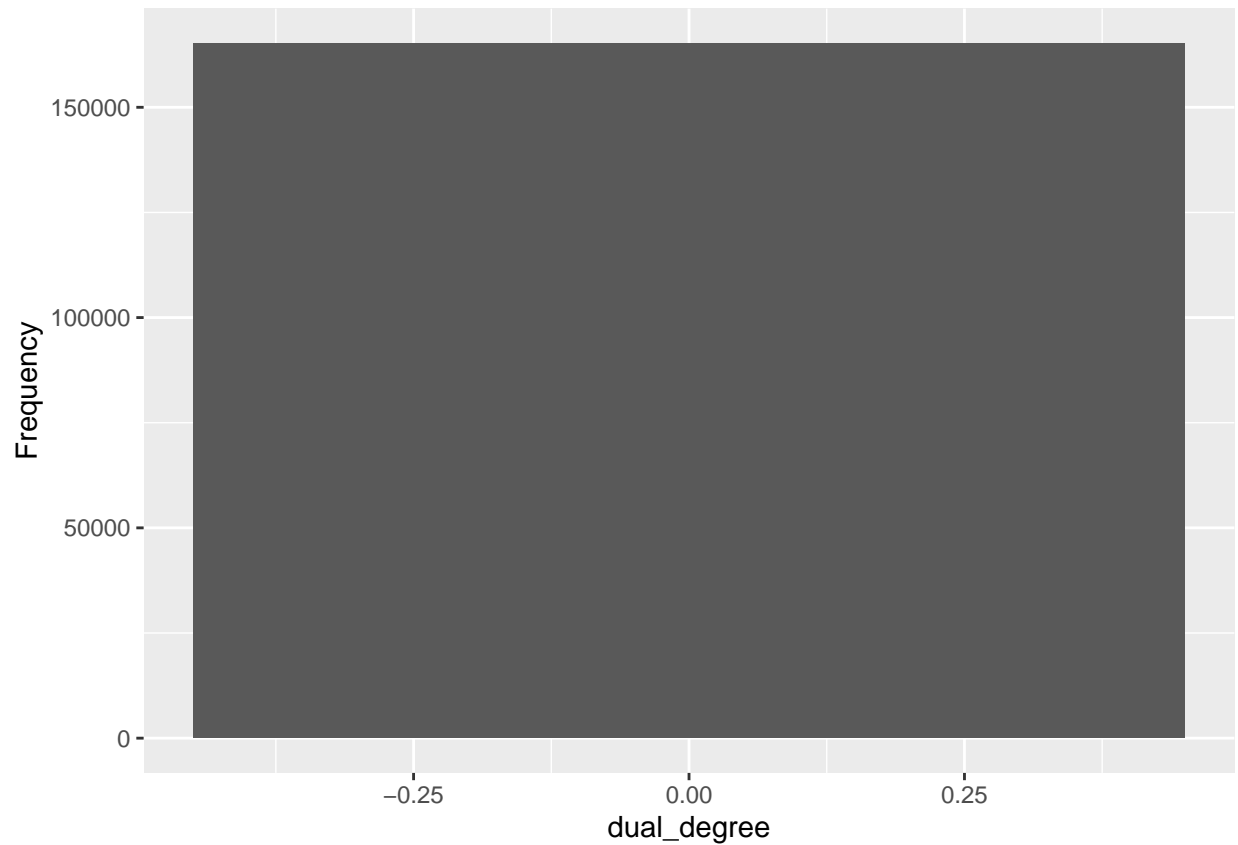
```
[1] -----  
[1] Variable: second_baccalaureate, type: numeric, requested: yes  
[1] Values (3 unique): NA, 0, 1  
[1] Missing: 99.7%
```

Warning: Removed 905409 rows containing non-finite values ('stat_count()').



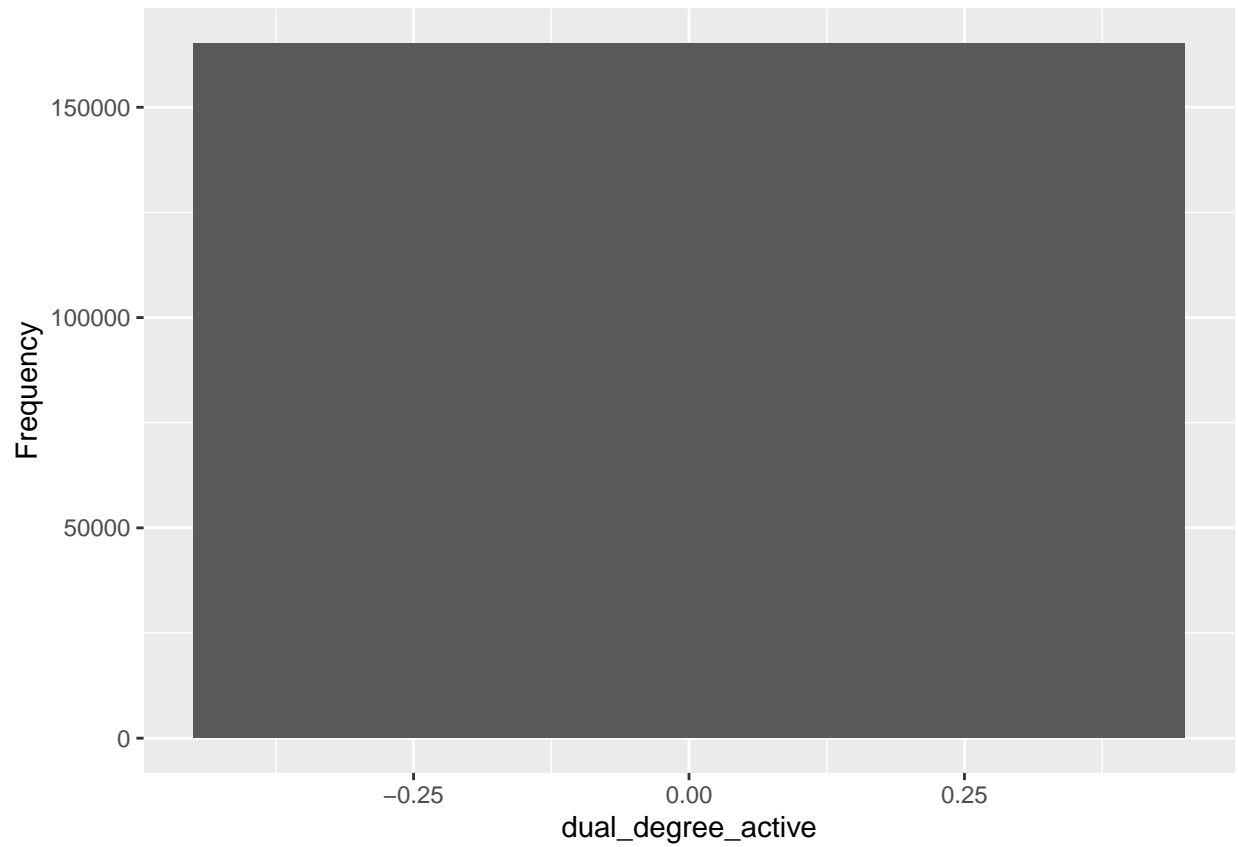
```
[1] -----
[1] Variable: dual_degree, type: numeric, requested: yes
[1] Values (2 unique): NA, 0
[1] Missing: 81.8%
  Group.1      x
1   2006 0.9990107
2   2007 0.9995634
3   2008 0.9995467
4   2009 0.9993081
5   2010 0.9993125
6   2011 0.9980670
7   2012 0.9972661
8   2013 0.9945803
9   2014 0.9771465
10  2015 0.8589683
11  2016 0.6647133
12  2017 0.4210091
13  2018 0.1200225
```

Warning: Removed 742630 rows containing non-finite values ('stat_count()').



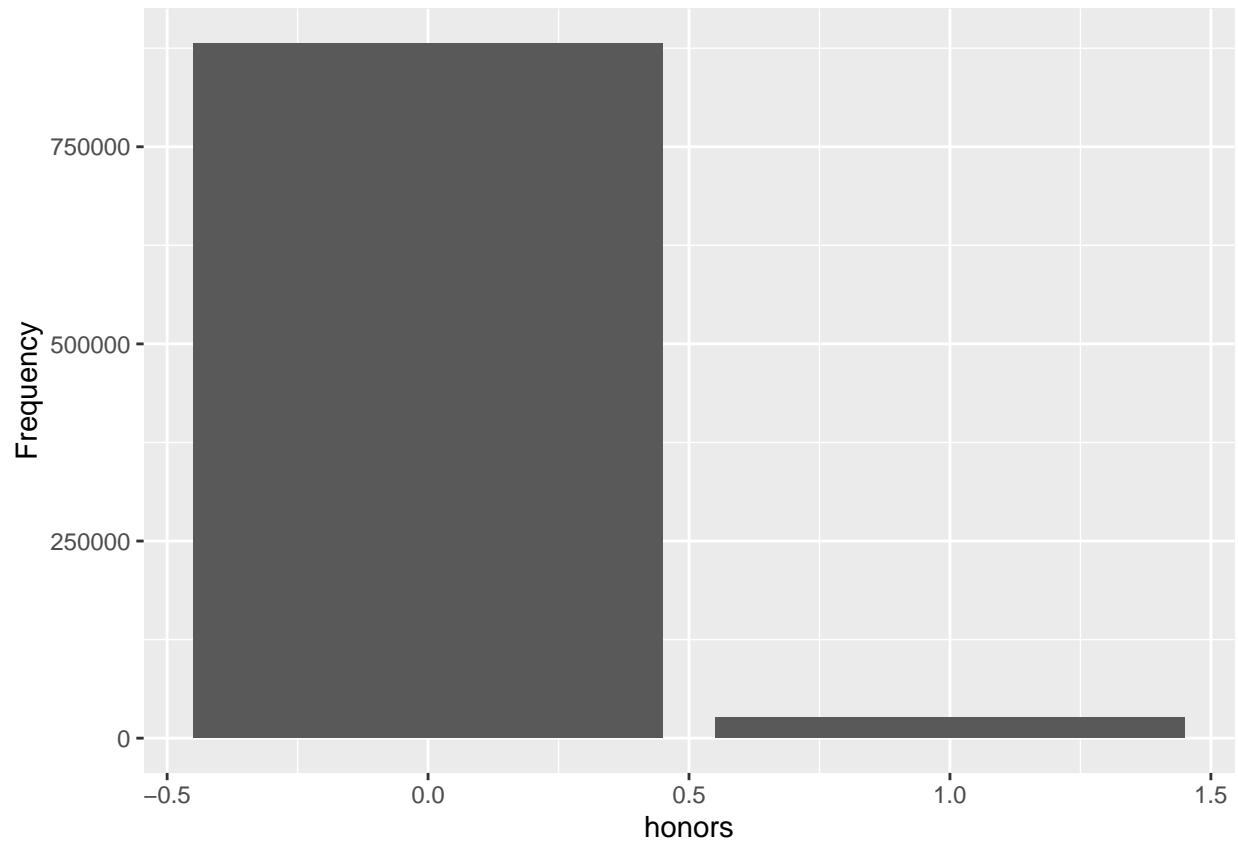
```
[1] -----
[1] Variable: dual_degree_active, type: numeric, requested: yes
[1] Values (2 unique): NA, 0
[1] Missing: 81.8%
  Group.1      x
1   2006 0.9990107
2   2007 0.9995634
3   2008 0.9995467
4   2009 0.9993081
5   2010 0.9993125
6   2011 0.9980670
7   2012 0.9972661
8   2013 0.9945803
9   2014 0.9771465
10  2015 0.8589683
11  2016 0.6647133
12  2017 0.4210091
13  2018 0.1200225
```

Warning: Removed 742630 rows containing non-finite values ('stat_count()').

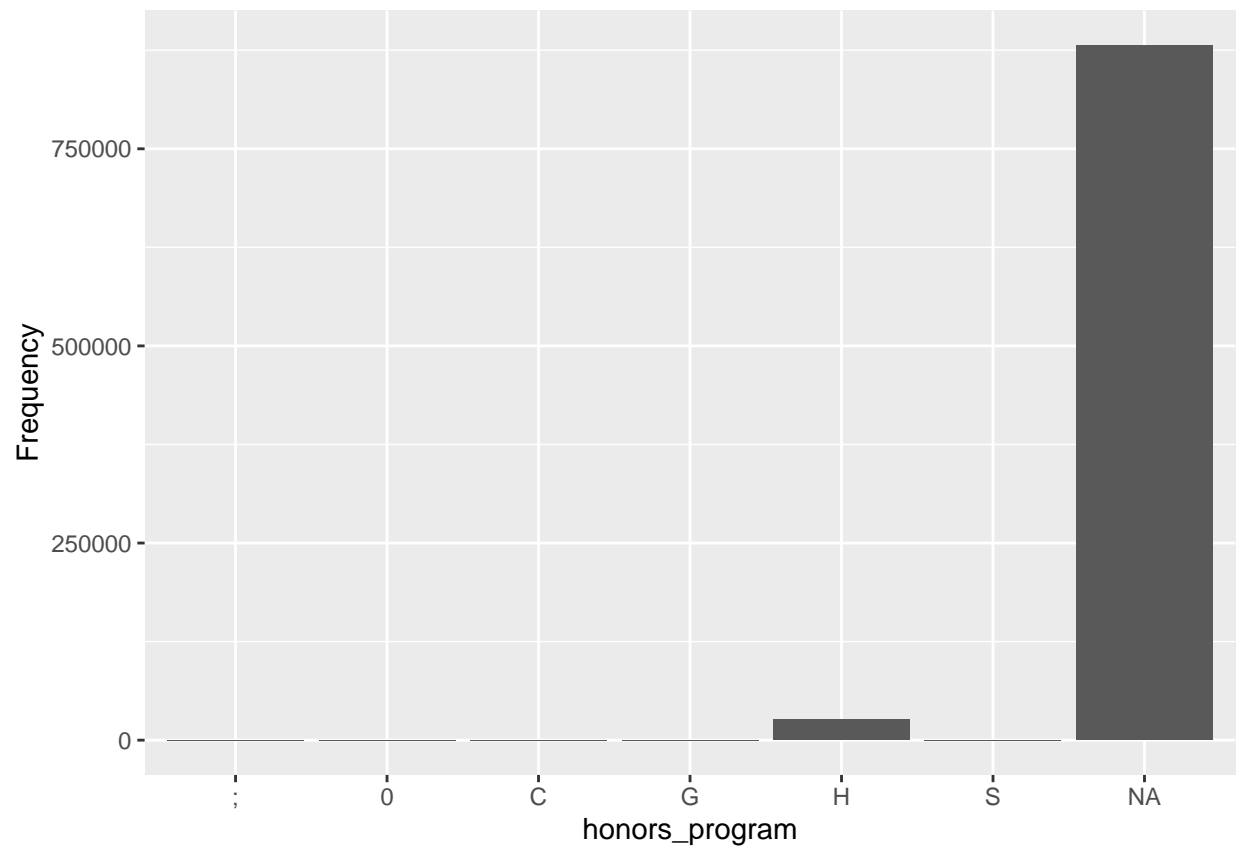


```
[1] -----  
[1] Variable: honors, type: numeric, requested: yes  
[1] Values (3 unique): 0, 1, NA  
[1] Missing: 0%
```

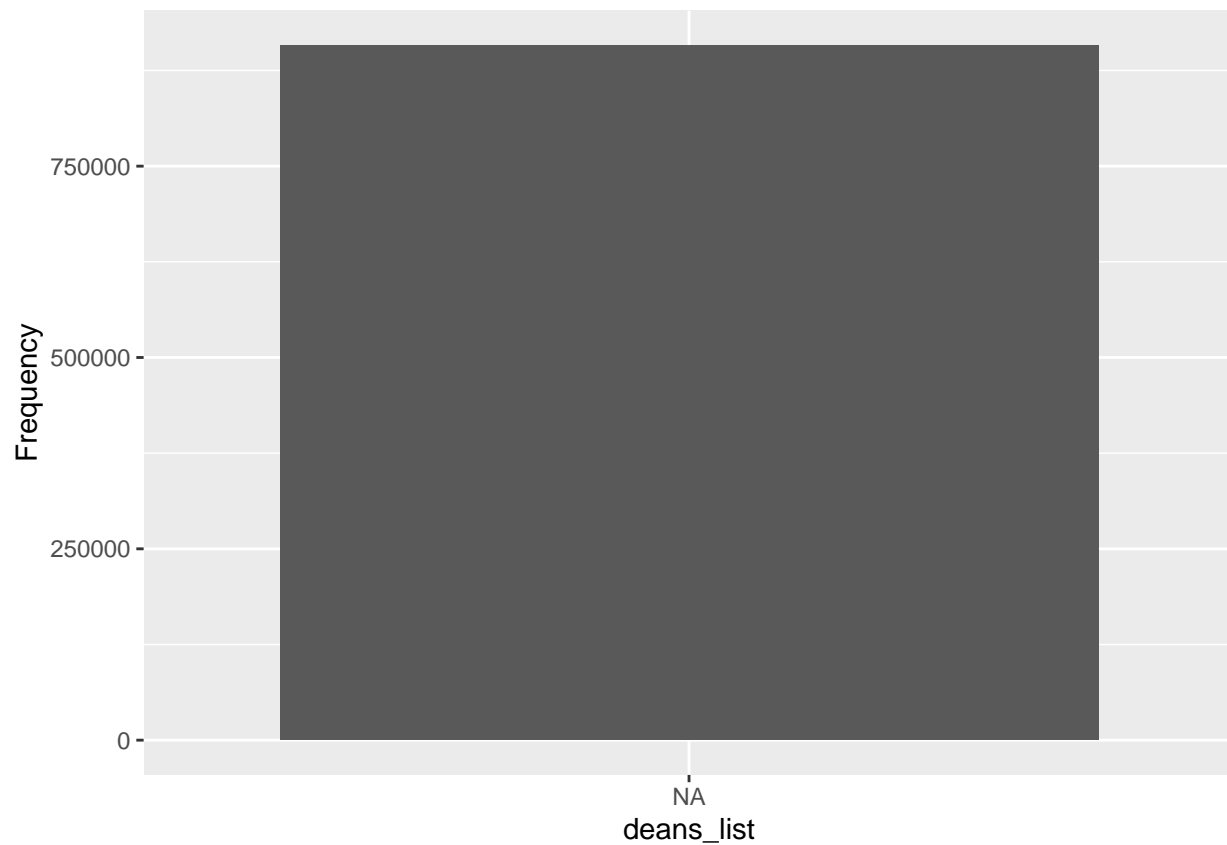
Warning: Removed 86 rows containing non-finite values ('stat_count()').



```
[1] -----  
[1] Variable: honors_program, type: character, requested: yes  
[1] Values (7 unique): NA, H, 0, ;, S, ...  
[1] Missing: 97.1%  
[1] Most missing: 2013 97.9%, Least missing: 2018 95.4%
```

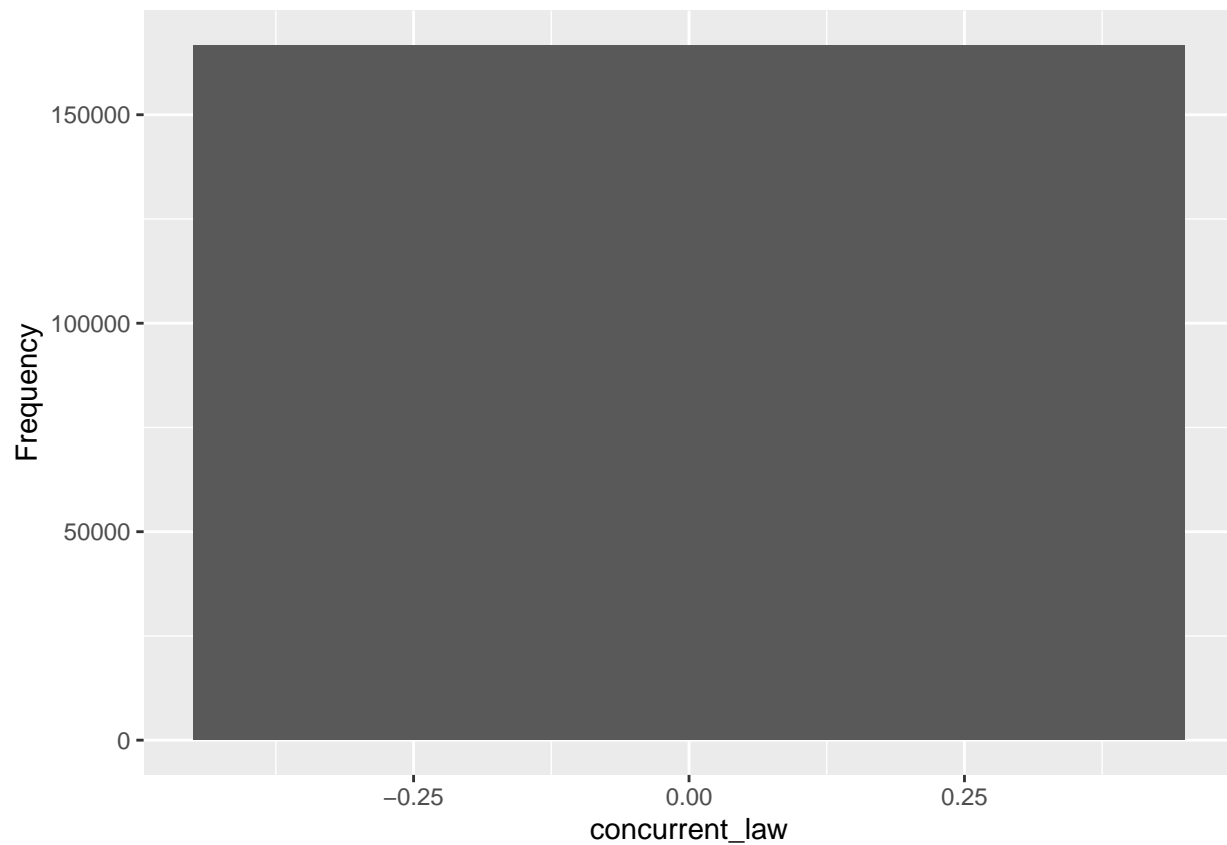


```
[1] -----  
[1] Variable: deans_list, type: logical, requested: yes  
[1] Values (1 unique): NA  
[1] Missing: 100%
```



```
[1] -----
[1] Variable: concurrent_low, type: numeric, requested: yes
[1] Values (2 unique): NA, 0
[1] Missing: 81.6%
  Group.1      x
1    2006 0.9989164
2    2007 0.9995310
3    2008 0.9995311
4    2009 0.9992562
5    2010 0.9993125
6    2011 0.9980093
7    2012 0.9972217
8    2013 0.9944993
9    2014 0.9769173
10   2015 0.8583636
11   2016 0.6640810
12   2017 0.4191856
13   2018 0.1058162
```

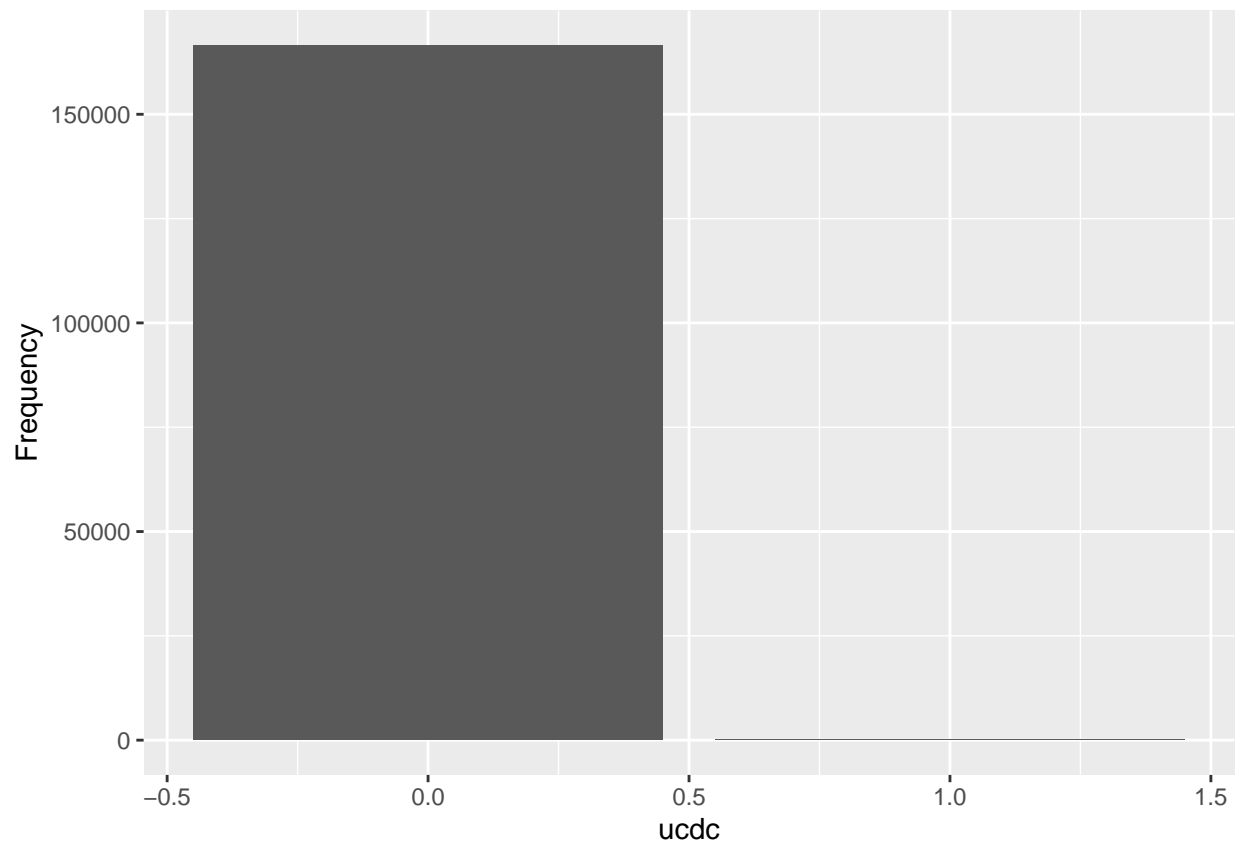
Warning: Removed 741188 rows containing non-finite values ('stat_count()').



```
[1] -----
[1] Variable: ucdc, type: numeric, requested: yes
[1] Values (3 unique): NA, 0, 1
[1] Missing: 81.6%
```

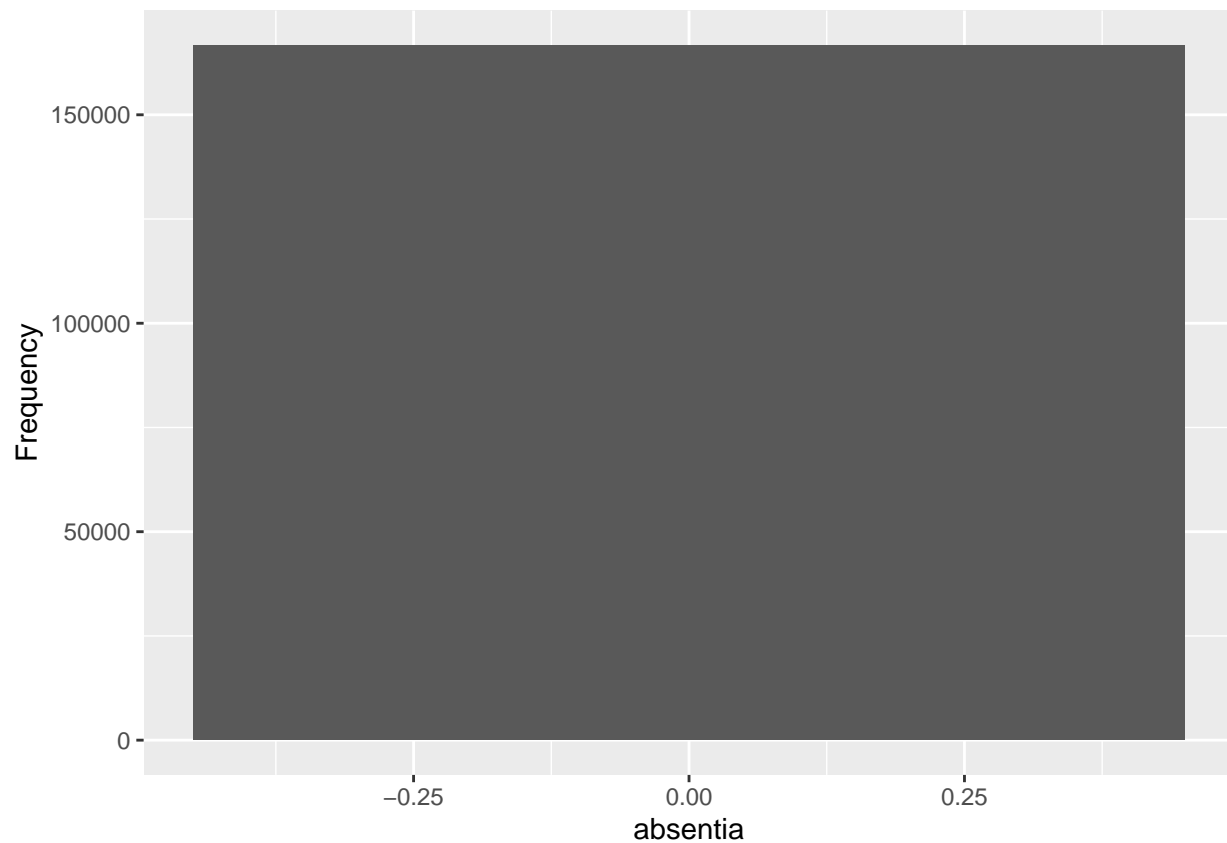
	Group.1	x
1	2006	0.9989164
2	2007	0.9995310
3	2008	0.9995311
4	2009	0.9992562
5	2010	0.9993125
6	2011	0.9980093
7	2012	0.9972217
8	2013	0.9944993
9	2014	0.9769173
10	2015	0.8583504
11	2016	0.6640472
12	2017	0.4191185
13	2018	0.1058037

Warning: Removed 741177 rows containing non-finite values ('stat_count()').



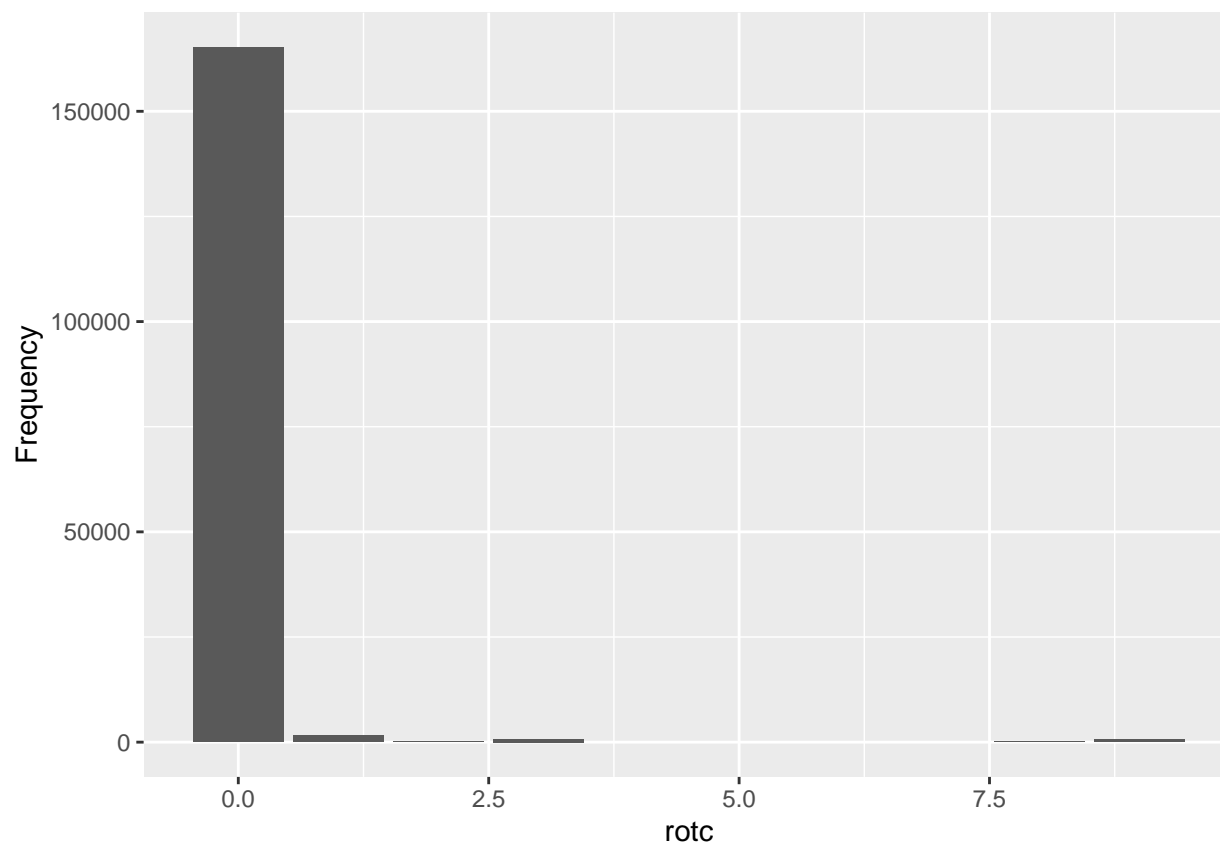
```
[1] -----
[1] Variable: absentia, type: numeric, requested: yes
[1] Values (2 unique): NA, 0
[1] Missing: 81.6%
  Group.1      x
1   2006 0.9989164
2   2007 0.9995310
3   2008 0.9995311
4   2009 0.9992562
5   2010 0.9993125
6   2011 0.9980093
7   2012 0.9972217
8   2013 0.9944993
9   2014 0.9769173
10  2015 0.8583504
11  2016 0.6640472
12  2017 0.4191185
13  2018 0.1058037
```

Warning: Removed 741177 rows containing non-finite values ('stat_count()').



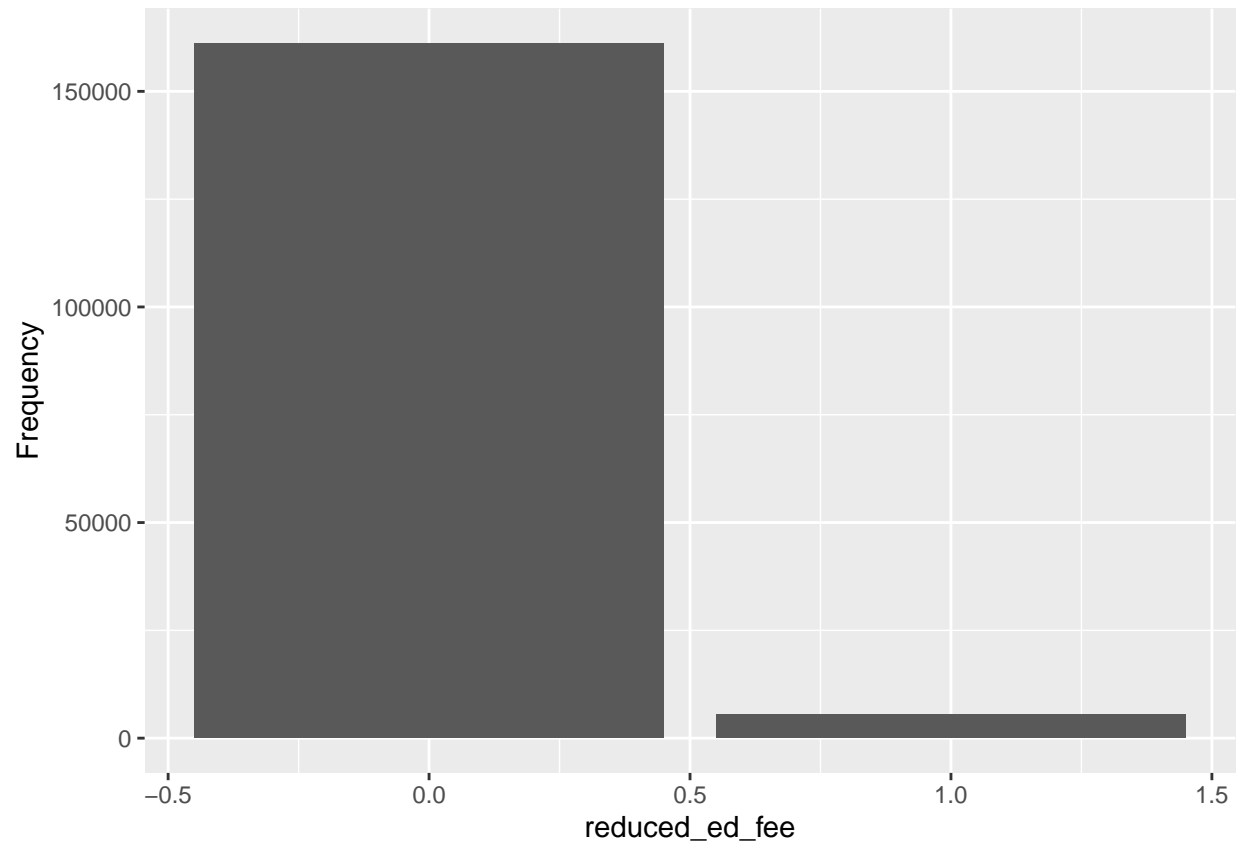
```
[1] -----
[1] Variable: rotc, type: numeric, requested: yes
[1] Values (7 unique): NA, 0, 1, 2, 9, ...
[1] Missing: 81.4%
  Group.1      x
1    2006 0.9990107
2    2007 0.9995634
3    2008 0.9995467
4    2009 0.9993081
5    2010 0.9993125
6    2011 0.9878107
7    2012 0.9905125
8    2013 0.9894445
9    2014 0.9734117
10   2015 0.8538676
11   2016 0.6572161
12   2017 0.4167692
13   2018 0.1190854
```

Warning: Removed 739299 rows containing non-finite values ('stat_count()').

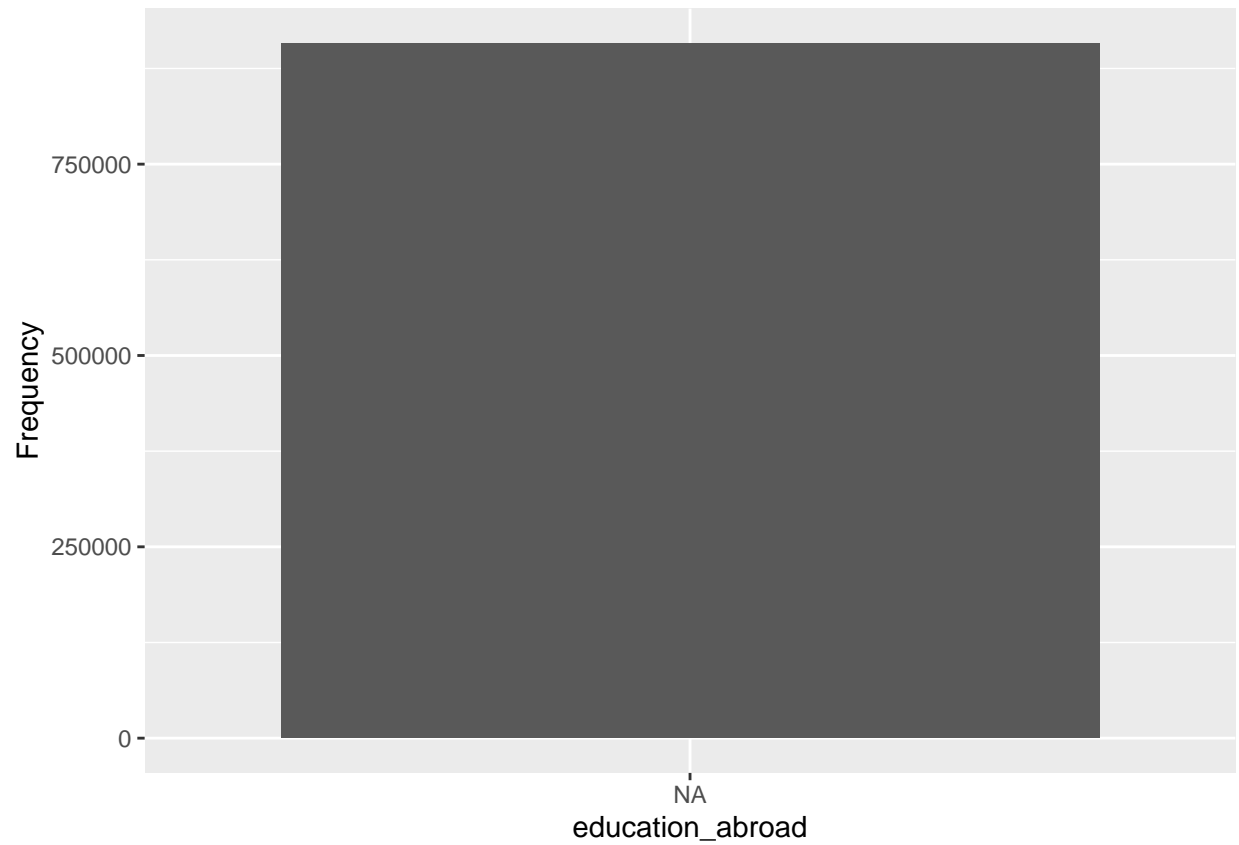


```
[1] -----
[1] Variable: reduced_ed_fee, type: numeric, requested: yes
[1] Values (3 unique): NA, 0, 1
[1] Missing: 81.6%
  Group.1      x
1    2006 0.9989164
2    2007 0.9995310
3    2008 0.9995311
4    2009 0.9992562
5    2010 0.9993125
6    2011 0.9980093
7    2012 0.9972217
8    2013 0.9944993
9    2014 0.9769173
10   2015 0.8583636
11   2016 0.6640472
12   2017 0.4191632
13   2018 0.1058662
```

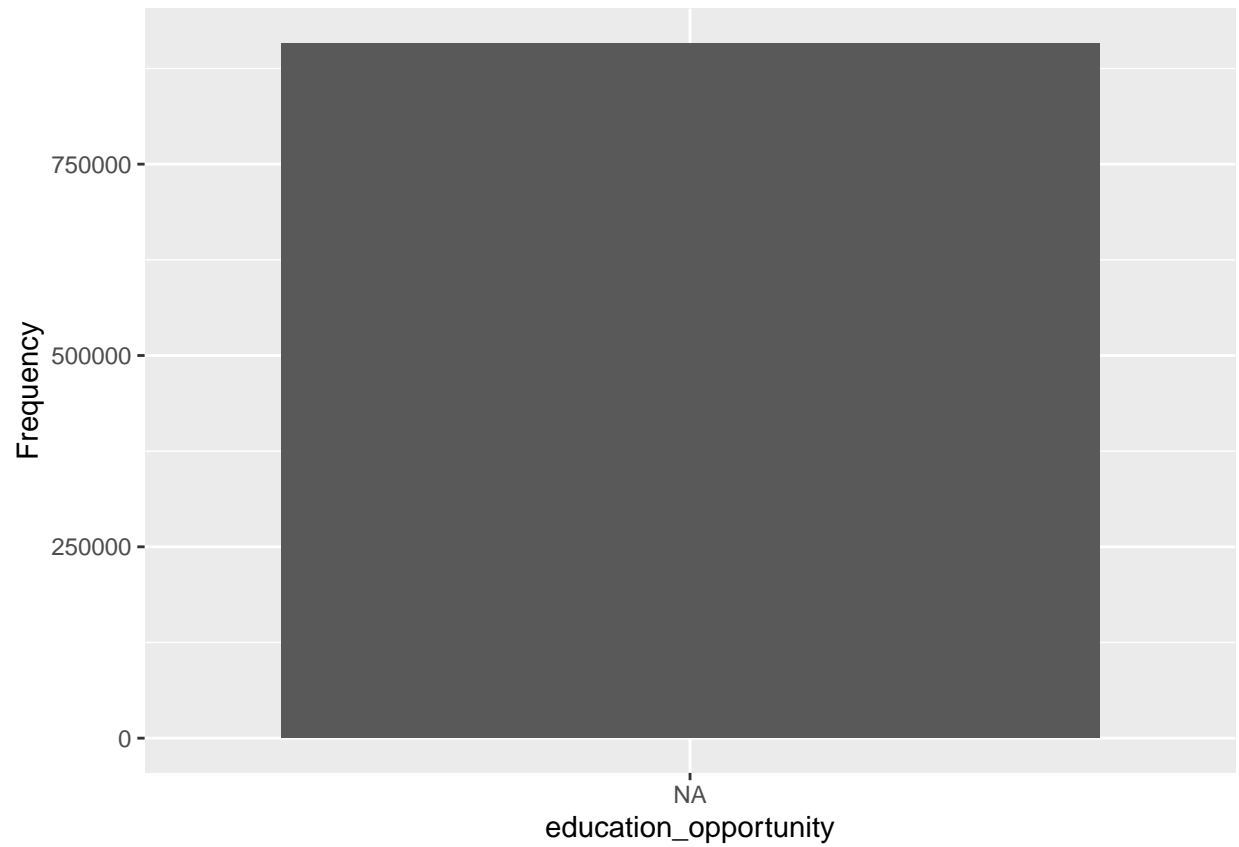
Warning: Removed 741187 rows containing non-finite values ('stat_count()').



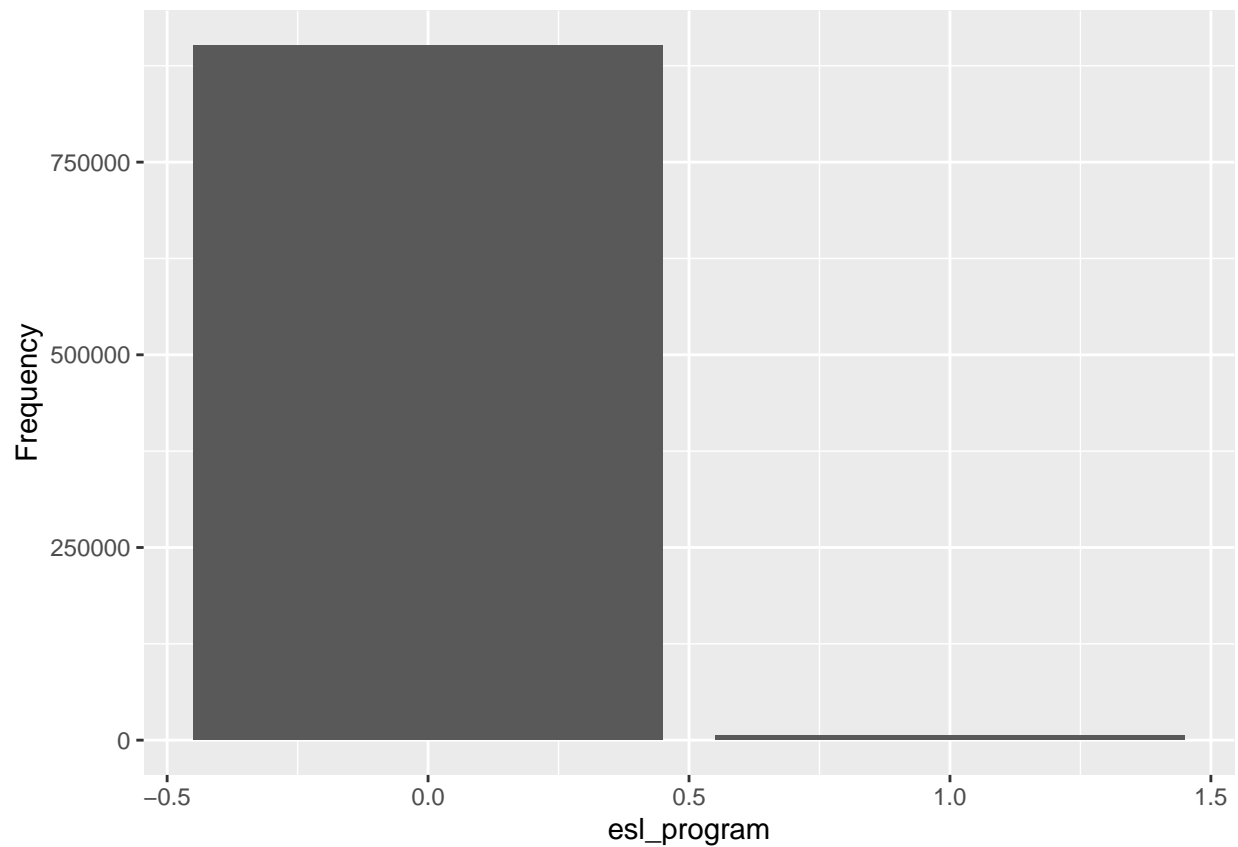
```
[1] -----  
[1] Variable: education_abroad, type: logical, requested: yes  
[1] Values (1 unique): NA  
[1] Missing: 100%
```



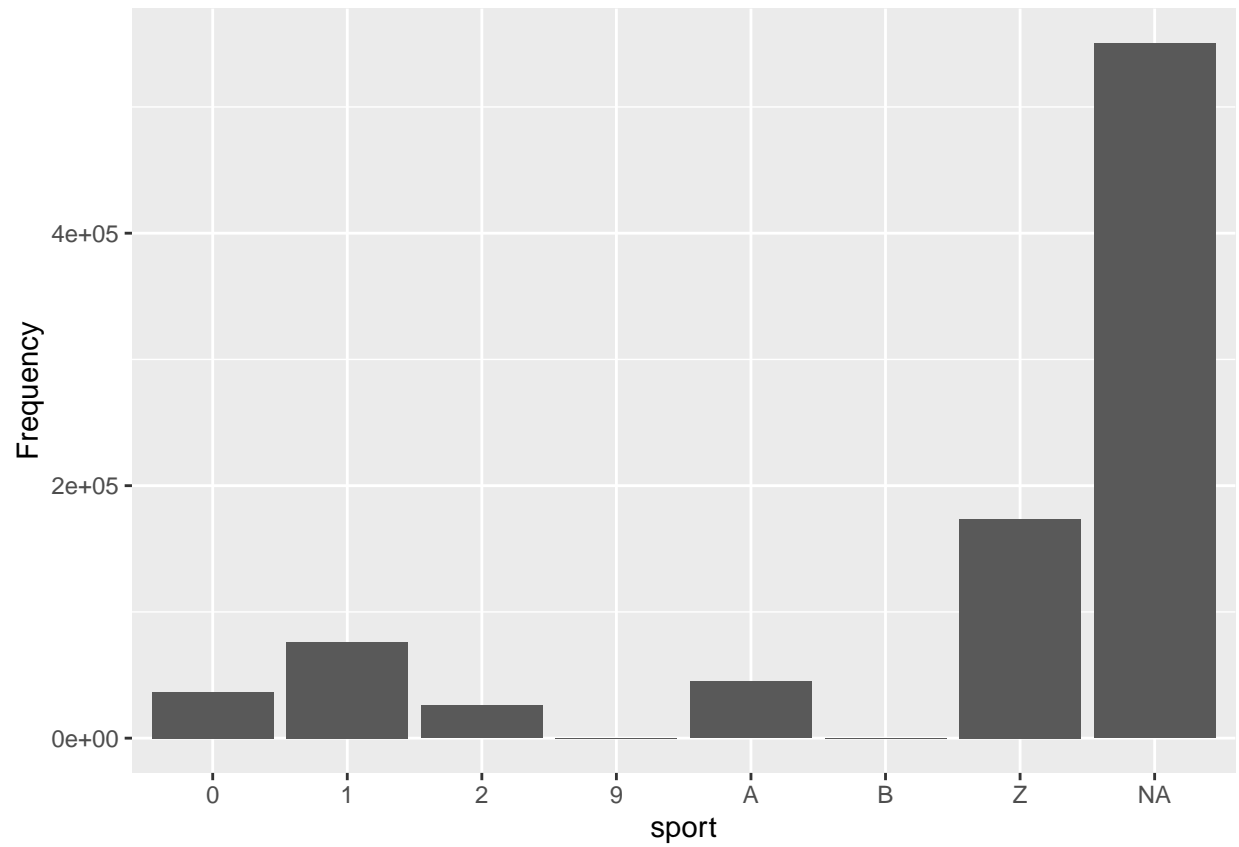
```
[1] -----  
[1] Variable: education_opportunity, type: logical, requested: yes  
[1] Values (1 unique): NA  
[1] Missing: 100%
```



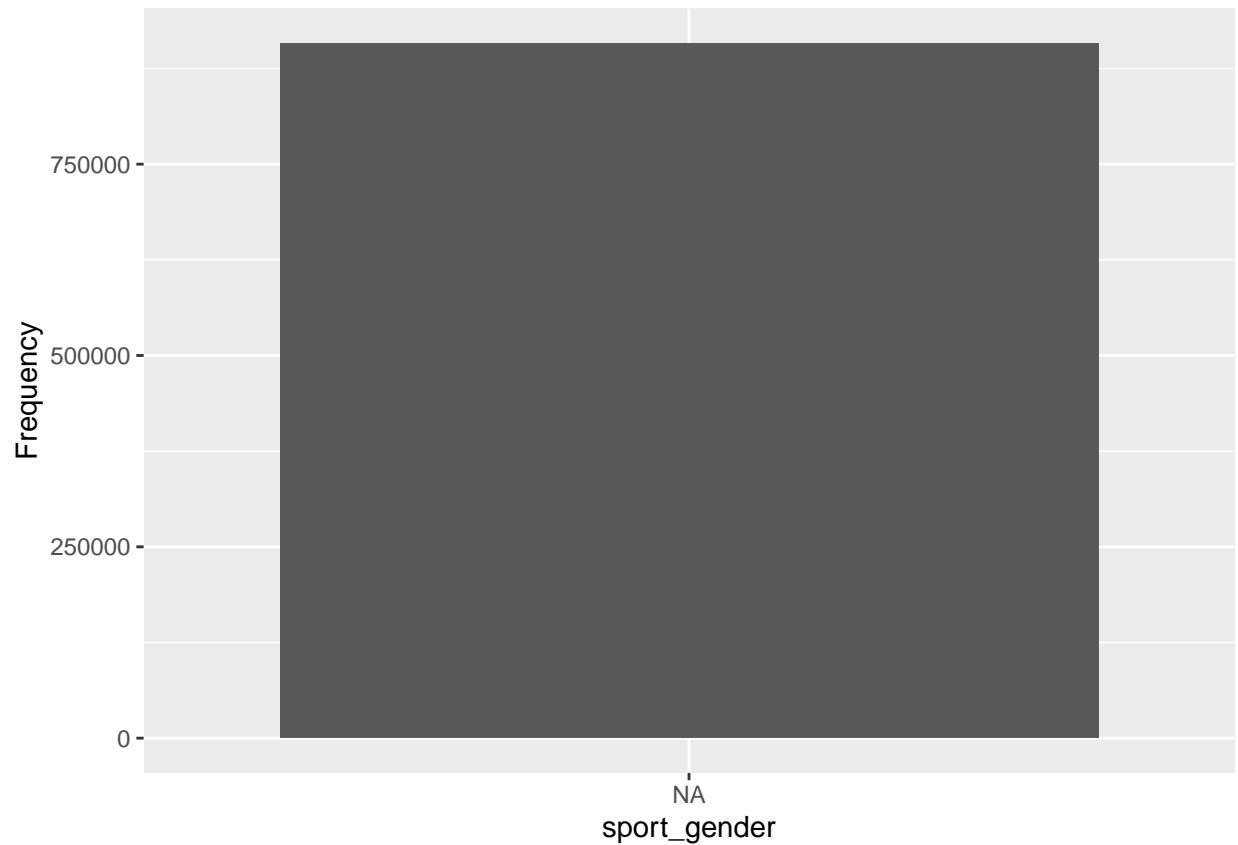
```
[1] -----  
[1] Variable: esl_program, type: numeric, requested: yes  
[1] Values (2 unique): 0, 1  
[1] Missing: 0%
```



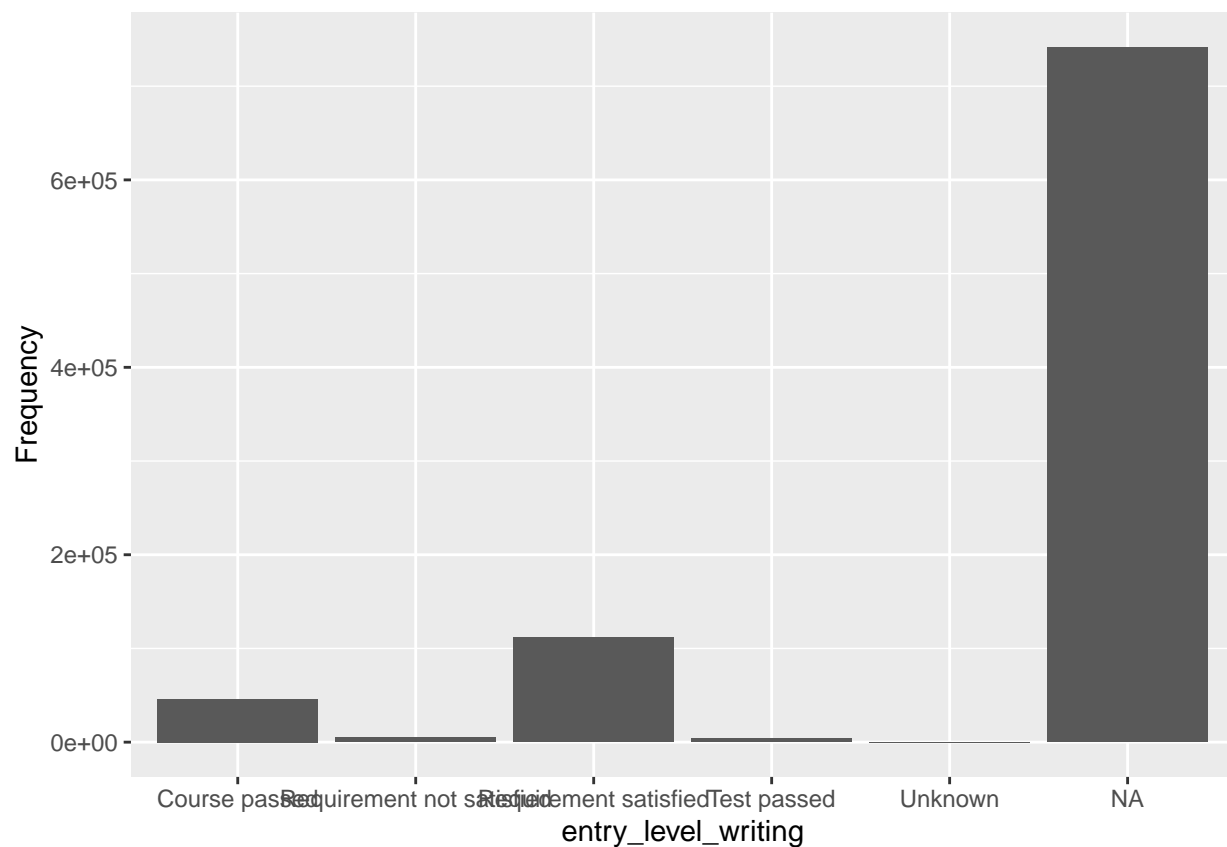
```
[1] -----
[1] Variable: sport, type: character, requested: yes
[1] Values (8 unique): 0, NA, Z, A, 1, ...
[1] Missing: 60.6%
  Group.1      x
1    2006 0.99870445
2    2007 0.99902974
3    2008 0.99738956
4    2009 0.99849515
5    2010 0.99825731
6    2011 0.08343551
7    2012 0.24816752
8    2013 0.31165022
9    2014 0.39688275
10   2015 0.44482568
11   2016 0.33672064
12   2017 0.58106052
13   2018 0.89420878
```



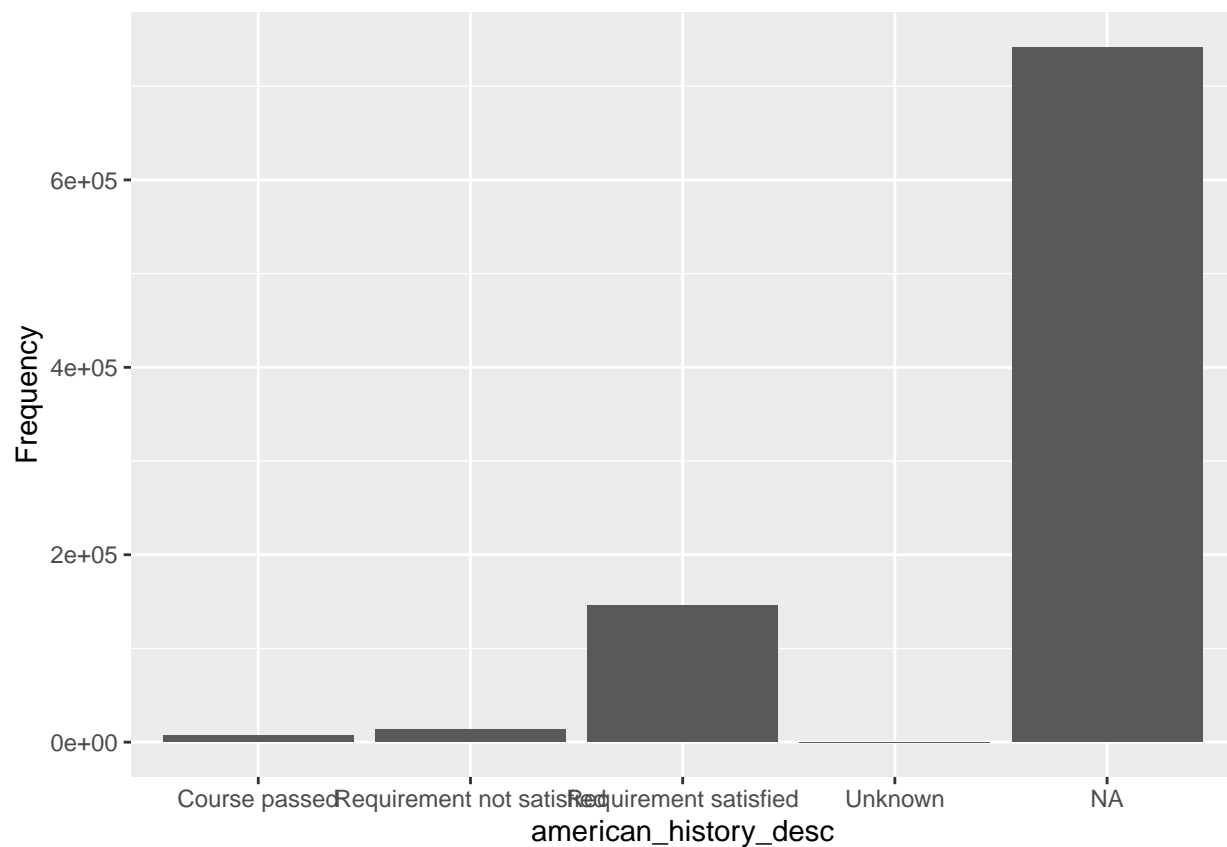
```
[1] -----
[1] Variable: sport_gender, type: logical, requested: yes
[1] Values (1 unique): NA
[1] Missing: 100%
```



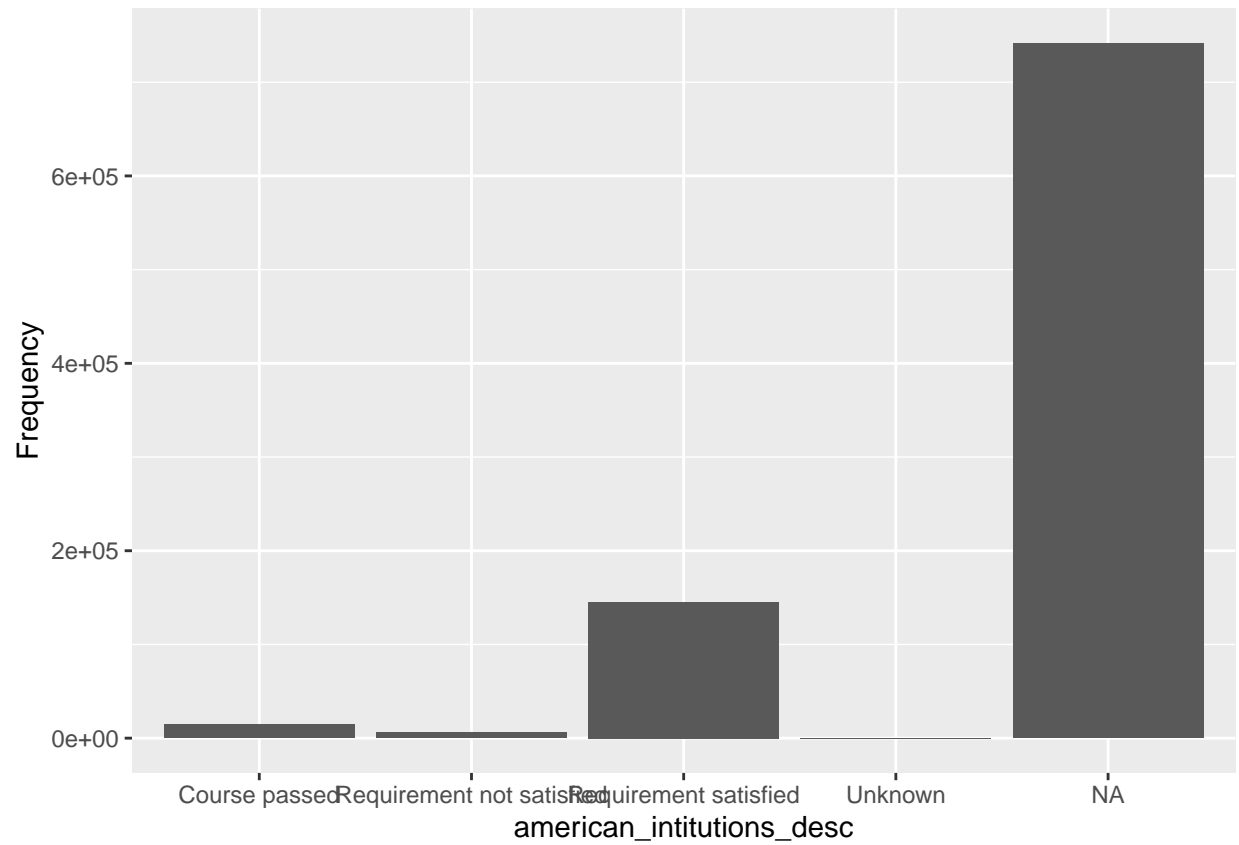
```
[1] -----
[1] Variable: entry_level_writing, type: character, requested: yes
[1] Values (6 unique): NA, Requirement satisfied, Course passed, Requirement not satisfied, Test passed
[1] Missing: 81.7%
  Group.1      x
1    2006 0.9989164
2    2007 0.9995310
3    2008 0.9995311
4    2009 0.9992562
5    2010 0.9993125
6    2011 0.9980093
7    2012 0.9972365
8    2013 0.9944993
9    2014 0.9769442
10   2015 0.8583504
11   2016 0.6641036
12   2017 0.4193310
13   2018 0.1087649
```



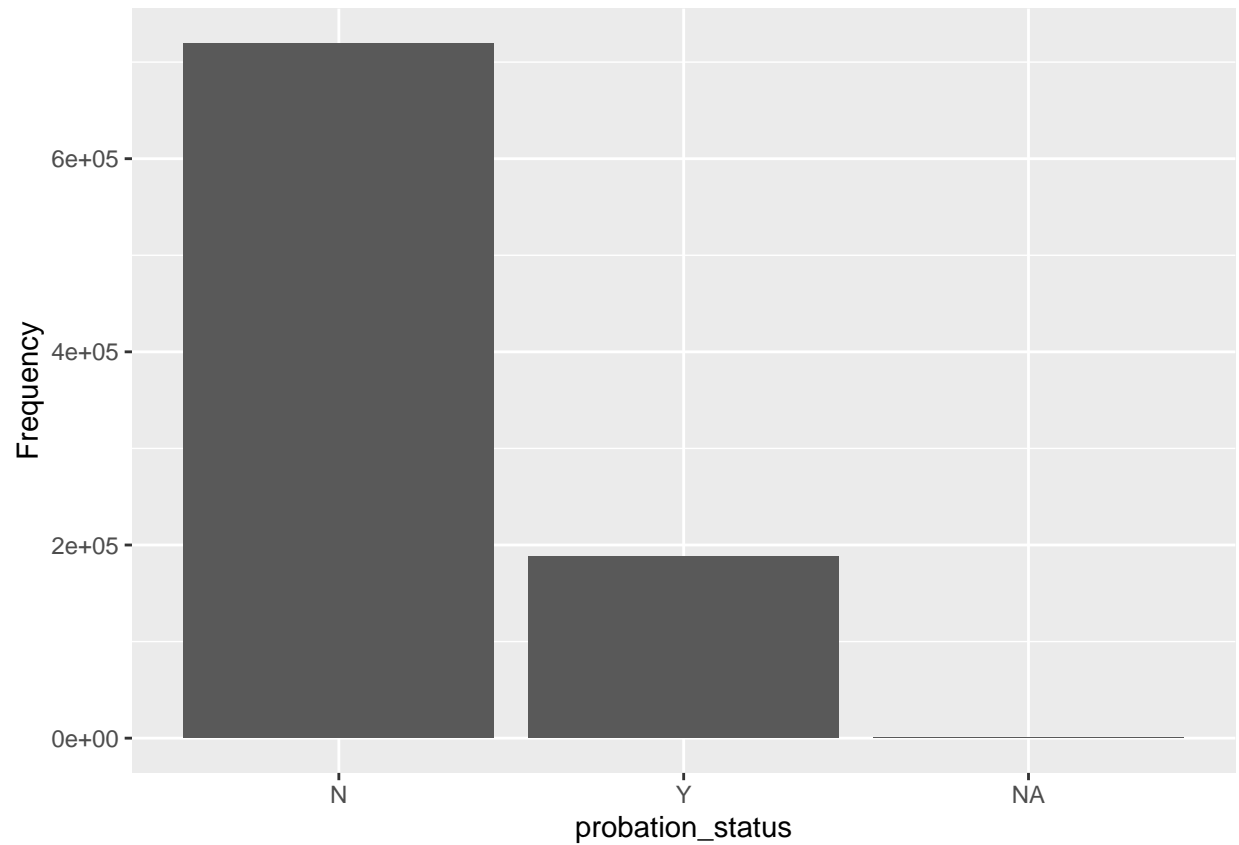
```
[1] -----
[1] Variable: american_history_desc, type: character, requested: yes
[1] Values (5 unique): NA, Requirement satisfied, Requirement not satisfied, Course passed, Unknown
[1] Missing: 81.7%
  Group.1      x
1    2006 0.9989164
2    2007 0.9995310
3    2008 0.9995311
4    2009 0.9992562
5    2010 0.9993125
6    2011 0.9980093
7    2012 0.9972365
8    2013 0.9944993
9    2014 0.9769442
10   2015 0.8583504
11   2016 0.6641375
12   2017 0.4195995
13   2018 0.1085400
```



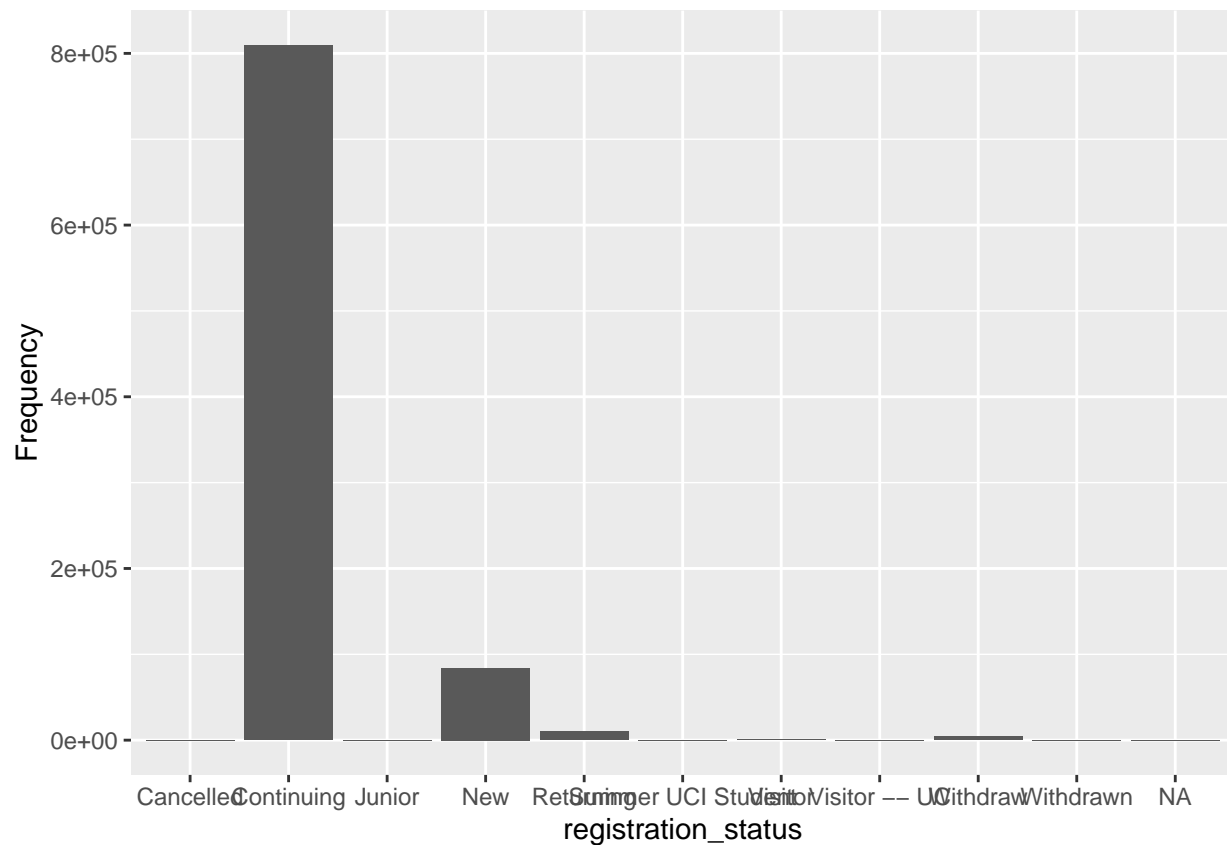
```
[1] -----
[1] Variable: american_intitutions_desc, type: character, requested: yes
[1] Values (5 unique): NA, Requirement satisfied, Course passed, Requirement not satisfied, Unknown
[1] Missing: 81.7%
  Group.1      x
1    2006 0.9989164
2    2007 0.9995310
3    2008 0.9995311
4    2009 0.9992562
5    2010 0.9993125
6    2011 0.9980093
7    2012 0.9972365
8    2013 0.9944993
9    2014 0.9769442
10   2015 0.8583504
11   2016 0.6641036
12   2017 0.4194765
13   2018 0.1088399
```

```
[1] -----
[1] Variable: probation_status, type: character, requested: yes
[1] Values (3 unique): N, Y, NA
[1] Missing: 0.1%
```

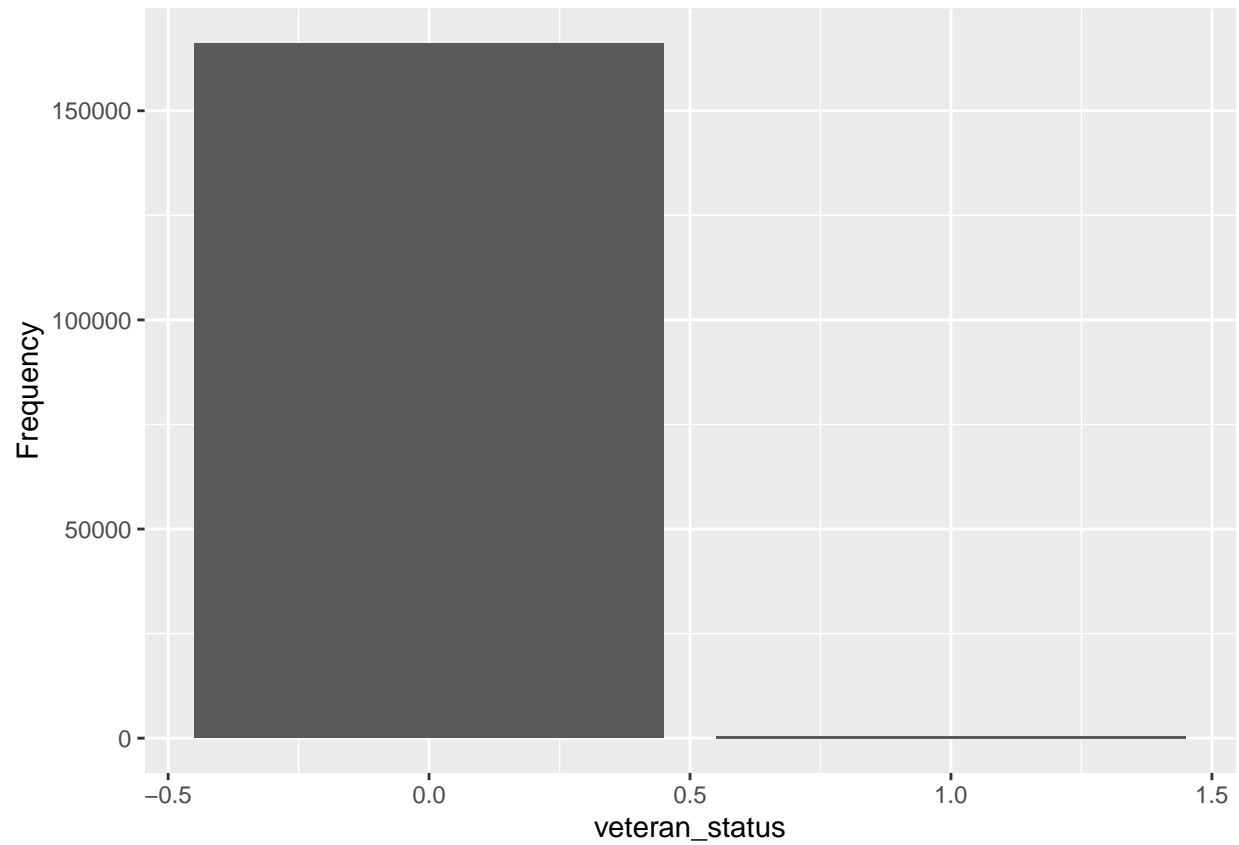


```
[1] -----  
[1] Variable: registration_status, type: character, requested: yes  
[1] Values (11 unique): New, Continuing, Returning, Withdraw, Visitor, ...  
[1] Missing: 0%
```

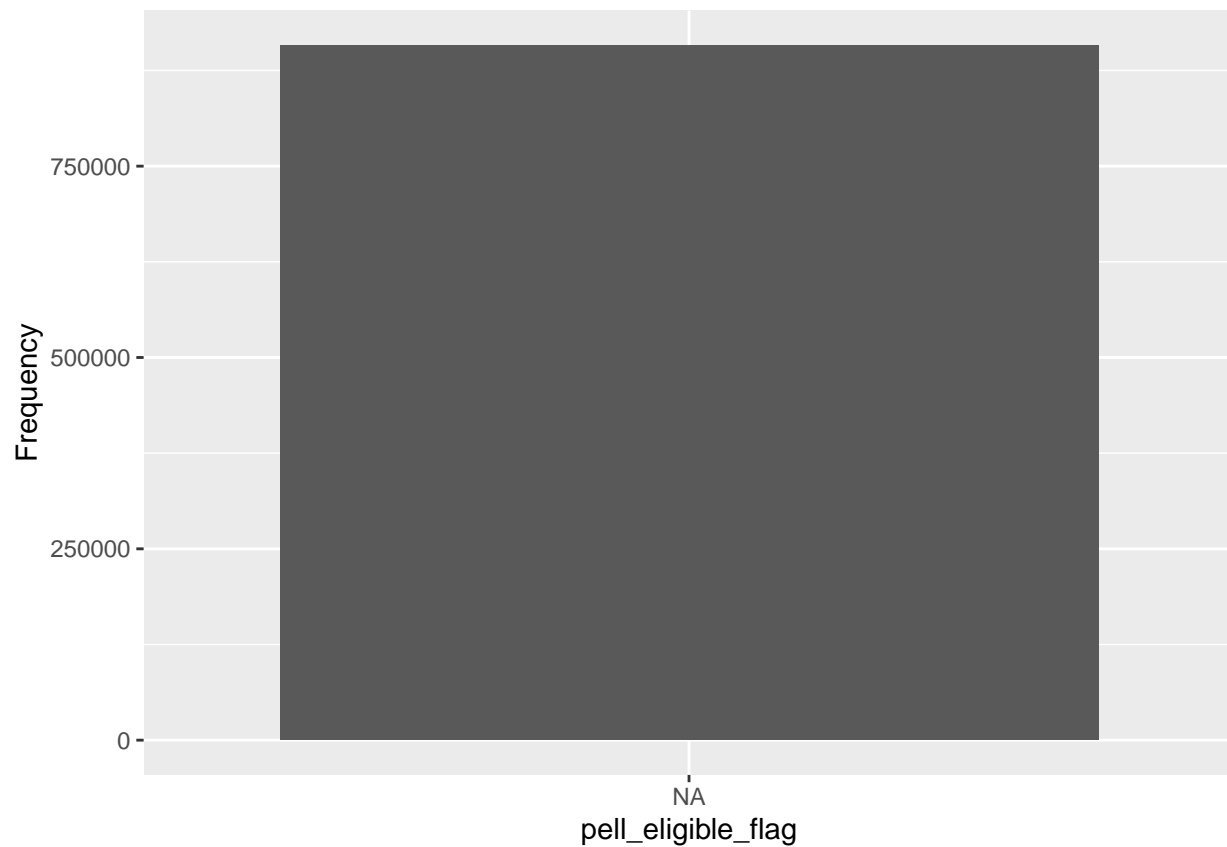


```
[1] -----
[1] Variable: veteran_status, type: numeric, requested: yes
[1] Values (3 unique): NA, 0, 1
[1] Missing: 81.6%
  Group.1      x
1    2006 0.9989164
2    2007 0.9995310
3    2008 0.9995311
4    2009 0.9992562
5    2010 0.9993125
6    2011 0.9980093
7    2012 0.9972217
8    2013 0.9944993
9    2014 0.9769173
10   2015 0.8583504
11   2016 0.6640472
12   2017 0.4191185
13   2018 0.1058037
```

Warning: Removed 741177 rows containing non-finite values ('stat_count()').

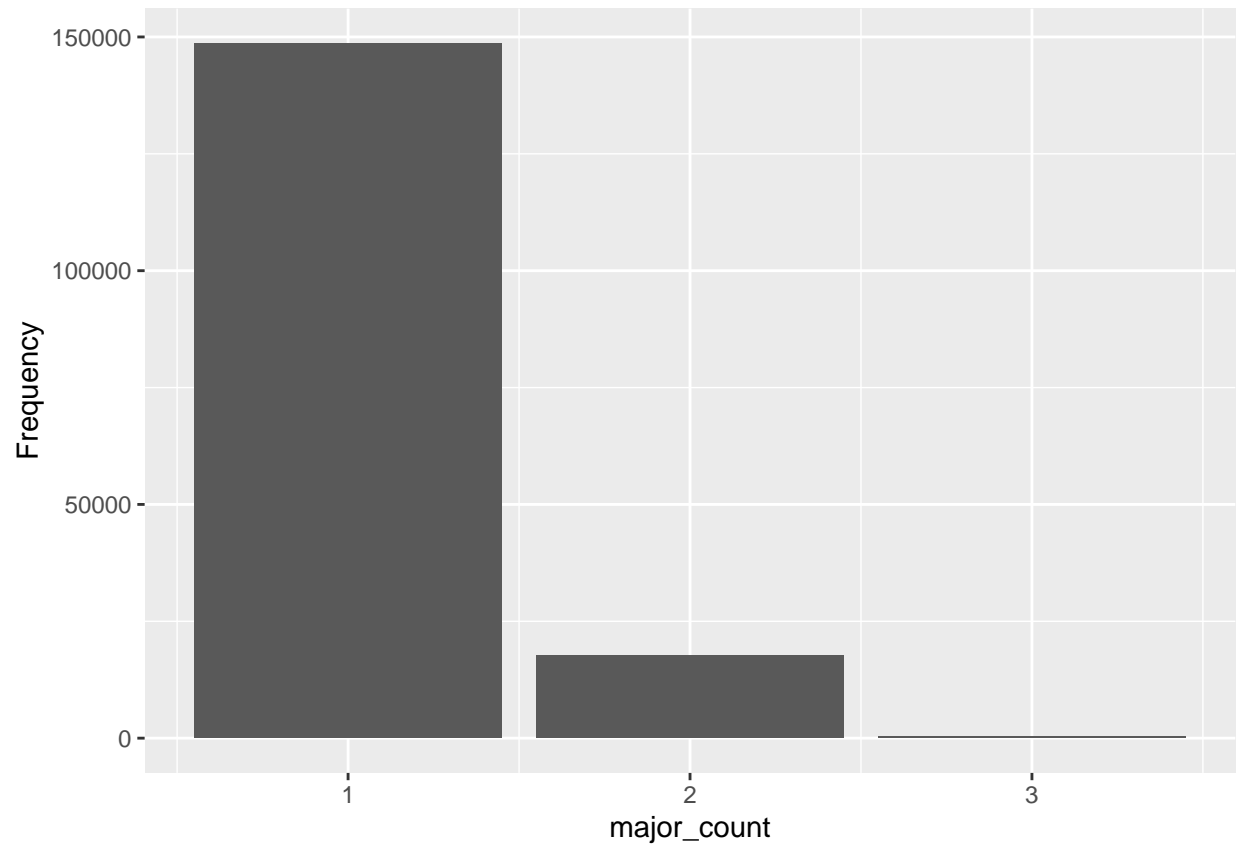


```
[1] -----  
[1] Variable: pell_eligible_flag, type: logical, requested: yes  
[1] Values (1 unique): NA  
[1] Missing: 100%
```

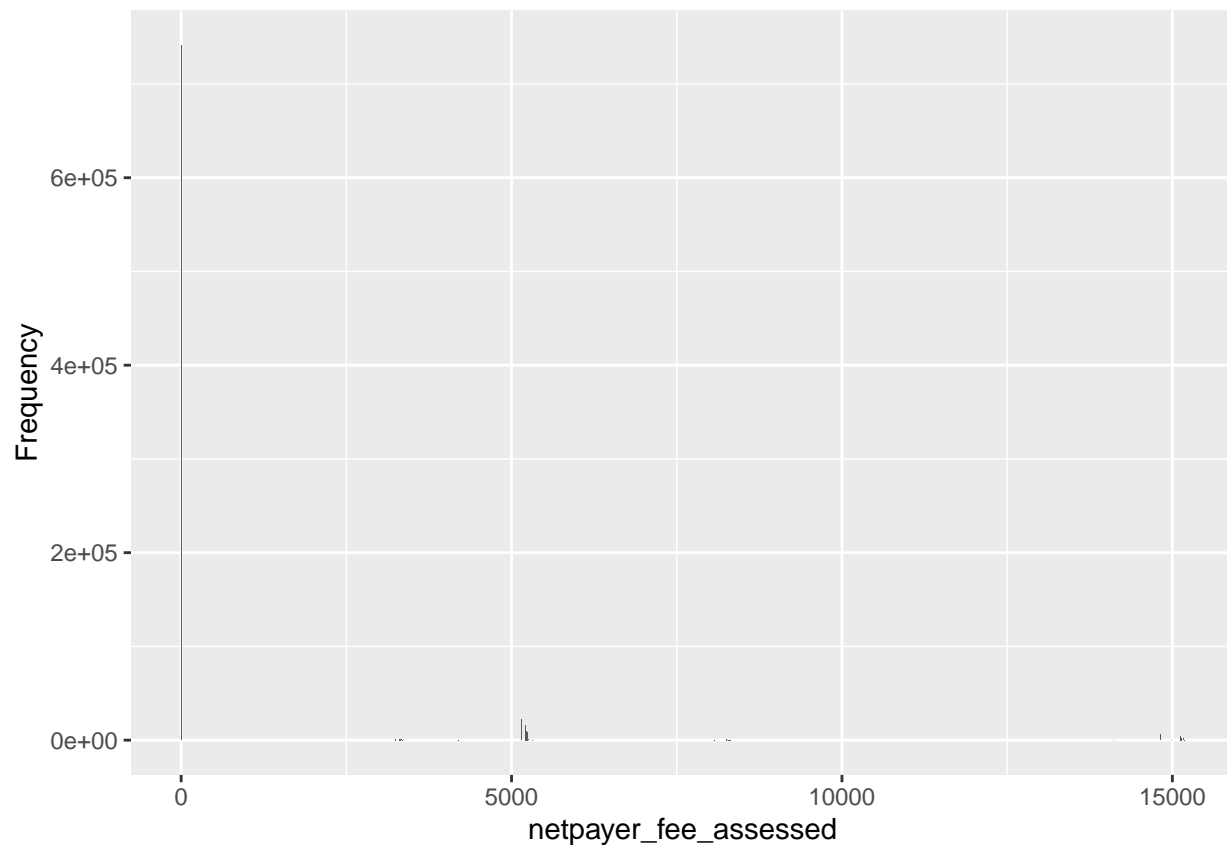


```
[1] -----
[1] Variable: major_count, type: numeric, requested: yes
[1] Values (4 unique): NA, 1, 2, 3
[1] Missing: 81.6%
  Group.1      x
1    2006 0.9989164
2    2007 0.9995310
3    2008 0.9995311
4    2009 0.9992562
5    2010 0.9993125
6    2011 0.9980093
7    2012 0.9972217
8    2013 0.9944993
9    2014 0.9769173
10   2015 0.8583504
11   2016 0.6640472
12   2017 0.4191185
13   2018 0.1058037
```

Warning: Removed 741177 rows containing non-finite values ('stat_count()').

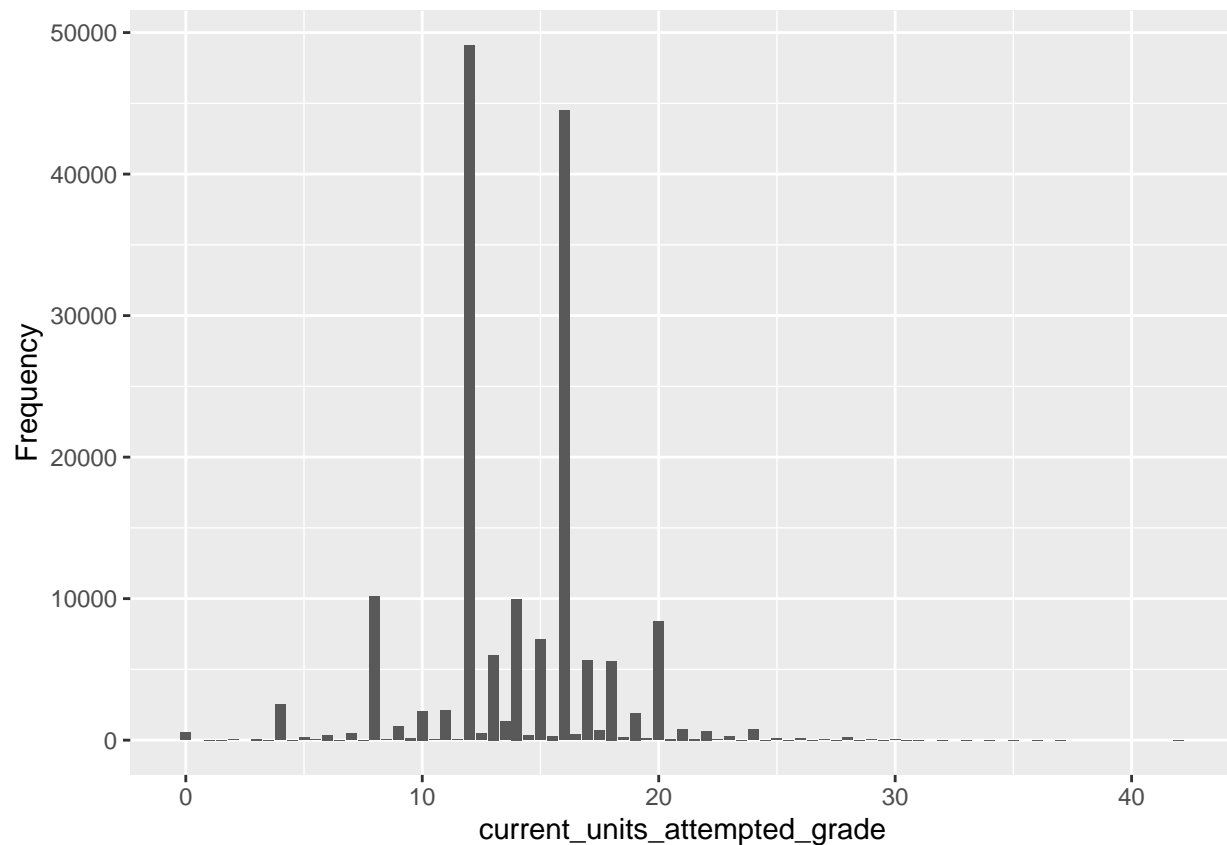


```
[1] -----  
[1] Variable: netpayer_fee_assessed, type: numeric, requested: yes  
[1] Values (56 unique): 0, 5207.04, 5207.0098, 5149.5298, 5245.75, ...  
[1] Missing: 0%
```



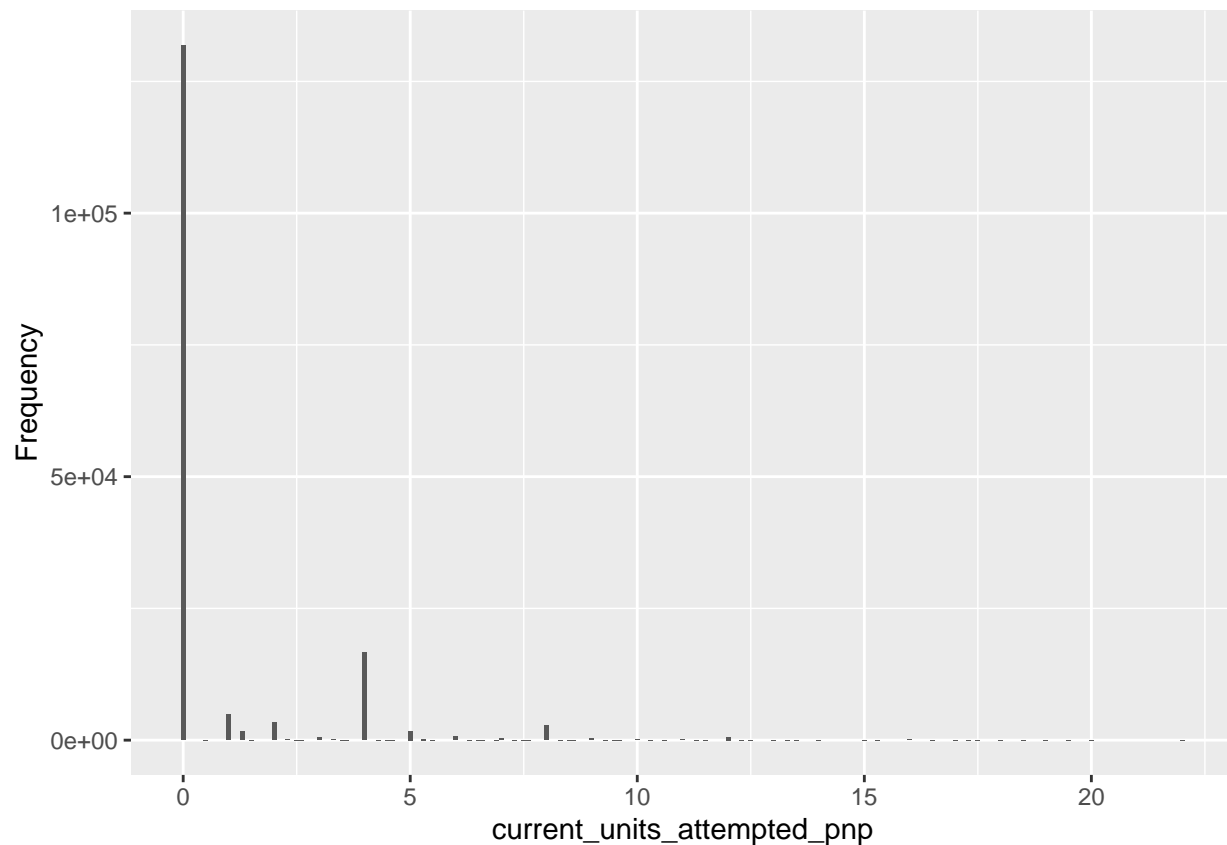
```
[1] -----
[1] Variable: current_units_attempted_grade, type: numeric, requested: yes
[1] Values (69 unique): NA, 12, 15, 8, 14, ...
[1] Missing: 81.8%
  Group.1      x
1    2006 0.9990107
2    2007 0.9995634
3    2008 0.9995467
4    2009 0.9993081
5    2010 0.9993125
6    2011 0.9980670
7    2012 0.9972661
8    2013 0.9945803
9    2014 0.9771465
10   2015 0.8589683
11   2016 0.6647133
12   2017 0.4210091
13   2018 0.1200225
```

Warning: Removed 742630 rows containing non-finite values ('stat_count()').



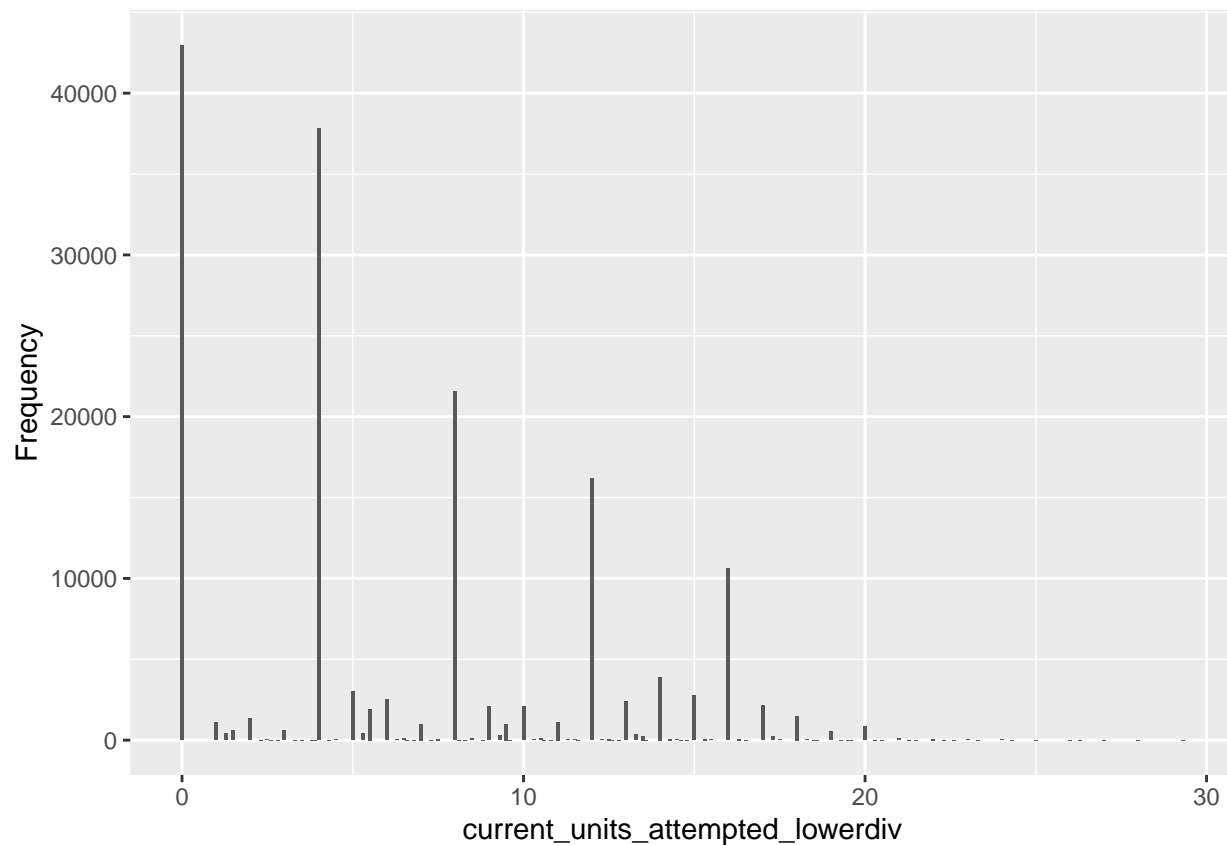
```
[1] -----
[1] Variable: current_units_attempted_pnp, type: numeric, requested: yes
[1] Values (64 unique): NA, 0, 6, 4, 2, ...
[1] Missing: 81.7%
  Group.1      x
1   2006 0.9989400
2   2007 0.9995472
3   2008 0.9995311
4   2009 0.9992735
5   2010 0.9993125
6   2011 0.9980237
7   2012 0.9972808
8   2013 0.9945128
9   2014 0.9770251
10  2015 0.8585082
11  2016 0.6642617
12  2017 0.4198792
13  2018 0.1080152
```

Warning: Removed 741470 rows containing non-finite values ('stat_count()').



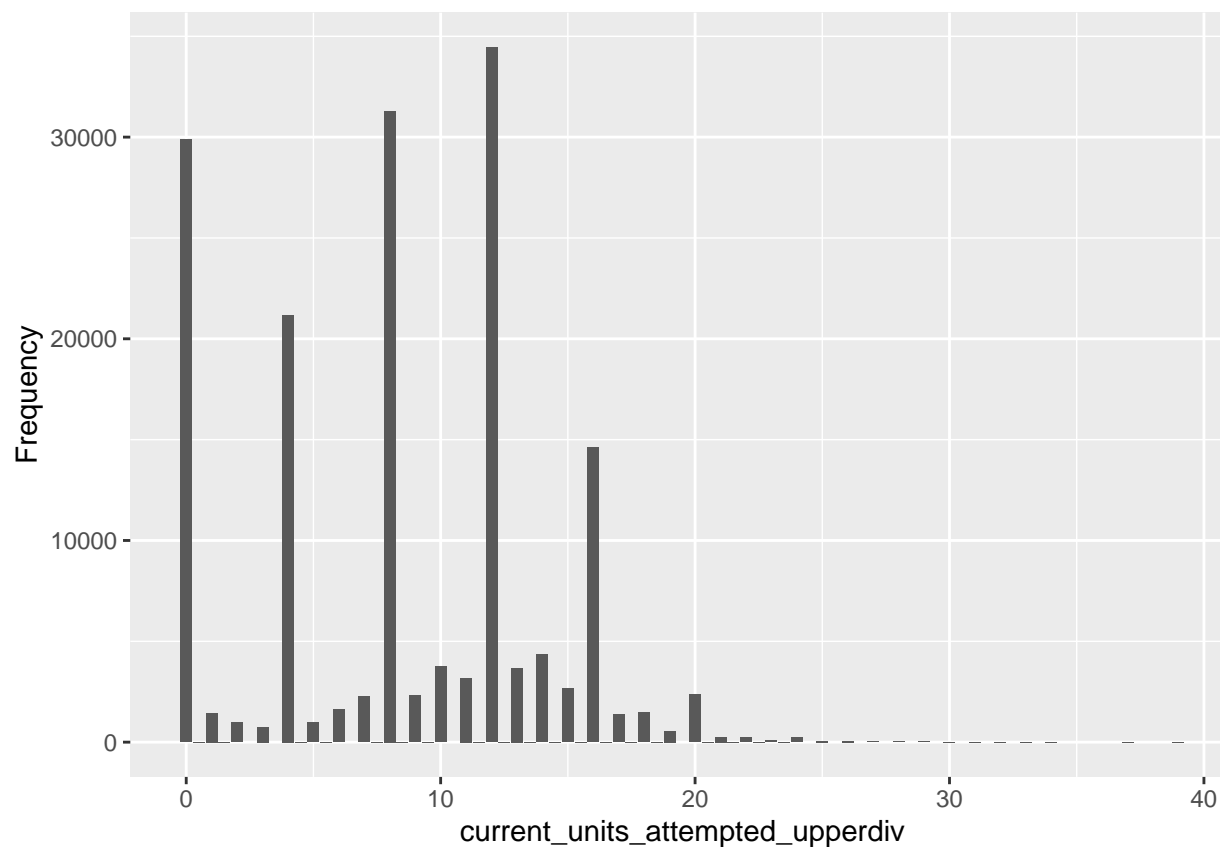
```
[1] -----
[1] Variable: current_units_attempted_lowerdiv, type: numeric, requested: yes
[1] Values (97 unique): NA, 8, 0, 12, 4, ...
[1] Missing: 81.8%
  Group.1      x
1    2006 0.9990578
2    2007 0.9995472
3    2008 0.9995780
4    2009 0.9992908
5    2010 0.9993445
6    2011 0.9980670
7    2012 0.9973547
8    2013 0.9946209
9    2014 0.9772678
10   2015 0.8591260
11   2016 0.6654021
12   2017 0.4232241
13   2018 0.1182857
```

Warning: Removed 742784 rows containing non-finite values ('stat_count()').



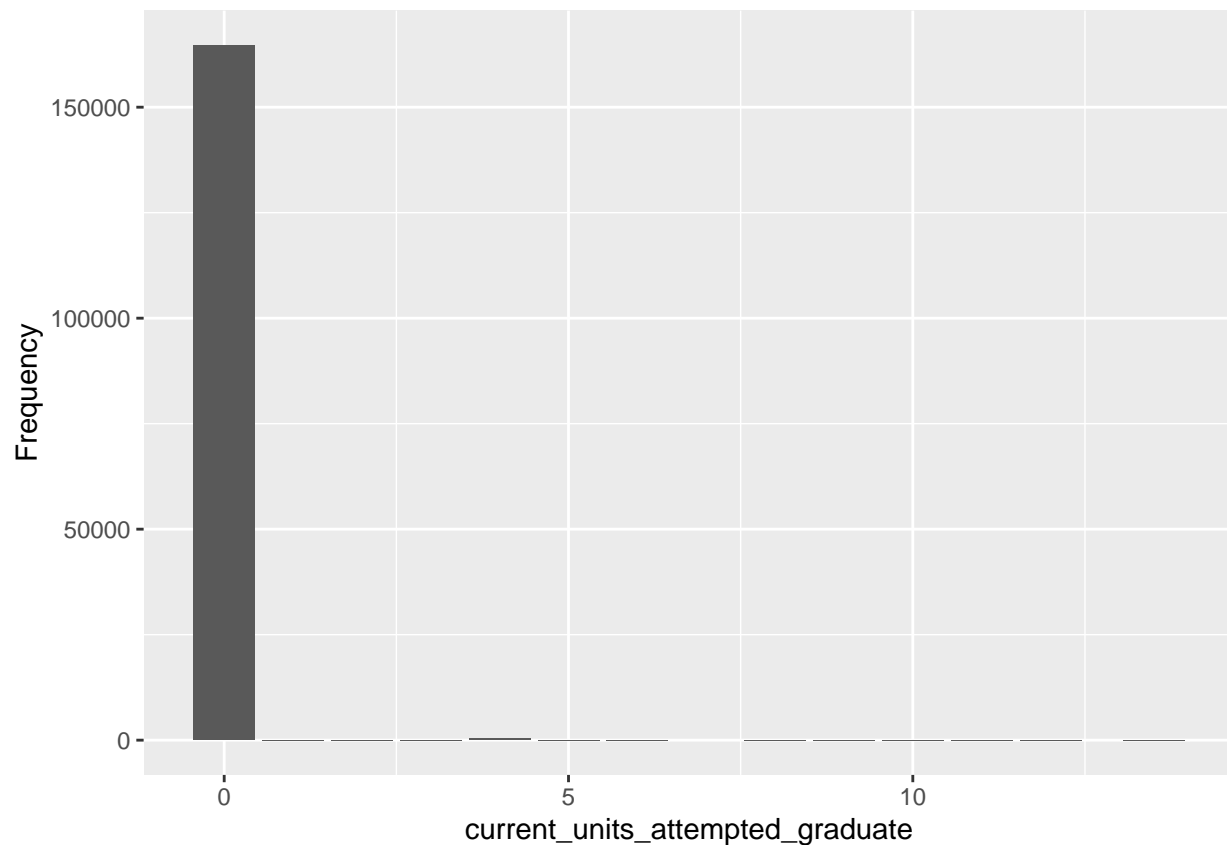
```
[1] -----
[1] Variable: current_units_attempted_upperdiv, type: numeric, requested: yes
[1] Values (57 unique): NA, 4, 15, 0, 12, ...
[1] Missing: 81.7%
  Group.1      x
1    2006 0.9989400
2    2007 0.9995472
3    2008 0.9995311
4    2009 0.9992735
5    2010 0.9993125
6    2011 0.9980237
7    2012 0.9972808
8    2013 0.9945263
9    2014 0.9770791
10   2015 0.8585871
11   2016 0.6643407
12   2017 0.4201365
13   2018 0.1093147
```

Warning: Removed 741615 rows containing non-finite values ('stat_count()').



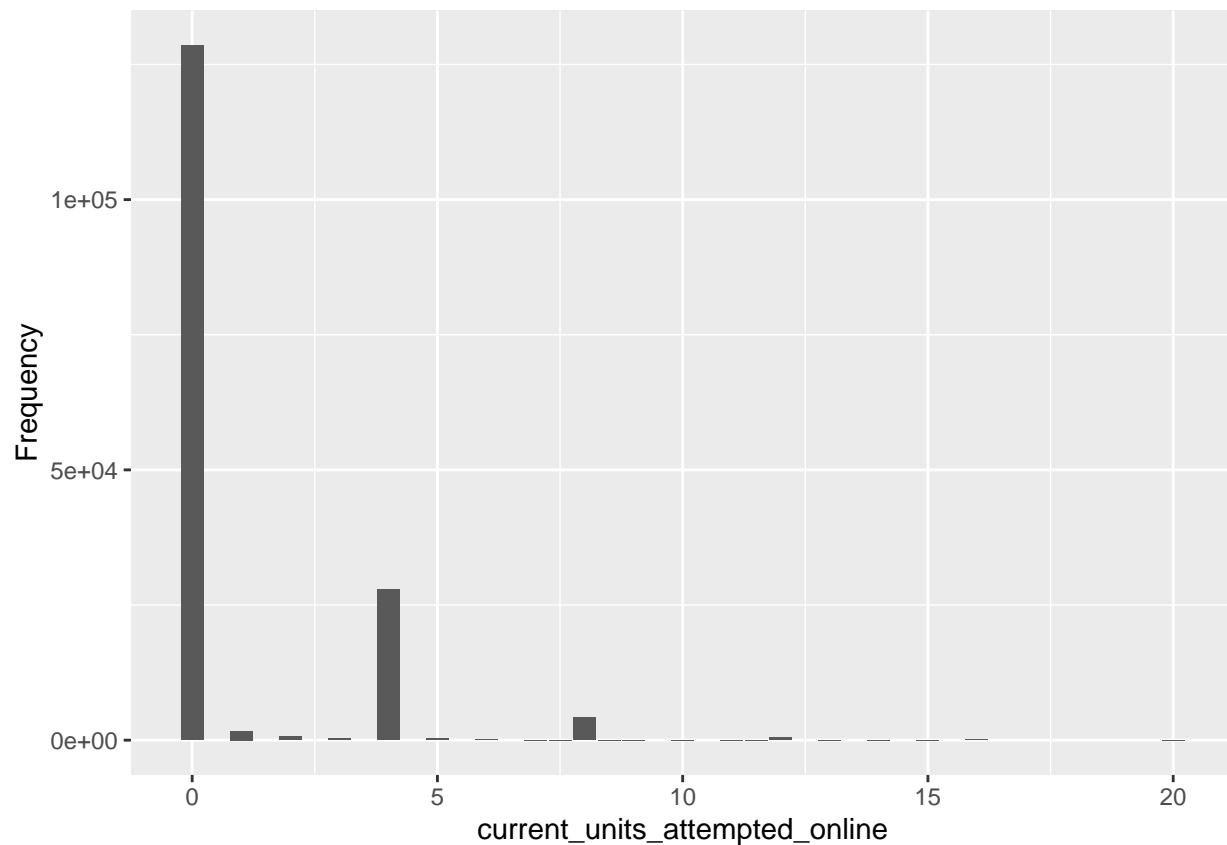
```
[1] -----
[1] Variable: current_units_attempted_graduate, type: numeric, requested: yes
[1] Values (14 unique): NA, 0, 2, 4, 8, ...
[1] Missing: 81.8%
  Group.1      x
1    2006 0.9990107
2    2007 0.9995634
3    2008 0.9995467
4    2009 0.9993081
5    2010 0.9993125
6    2011 0.9980670
7    2012 0.9972661
8    2013 0.9945803
9    2014 0.9771465
10   2015 0.8589683
11   2016 0.6647133
12   2017 0.4210091
13   2018 0.1200225
```

Warning: Removed 742630 rows containing non-finite values ('stat_count()').



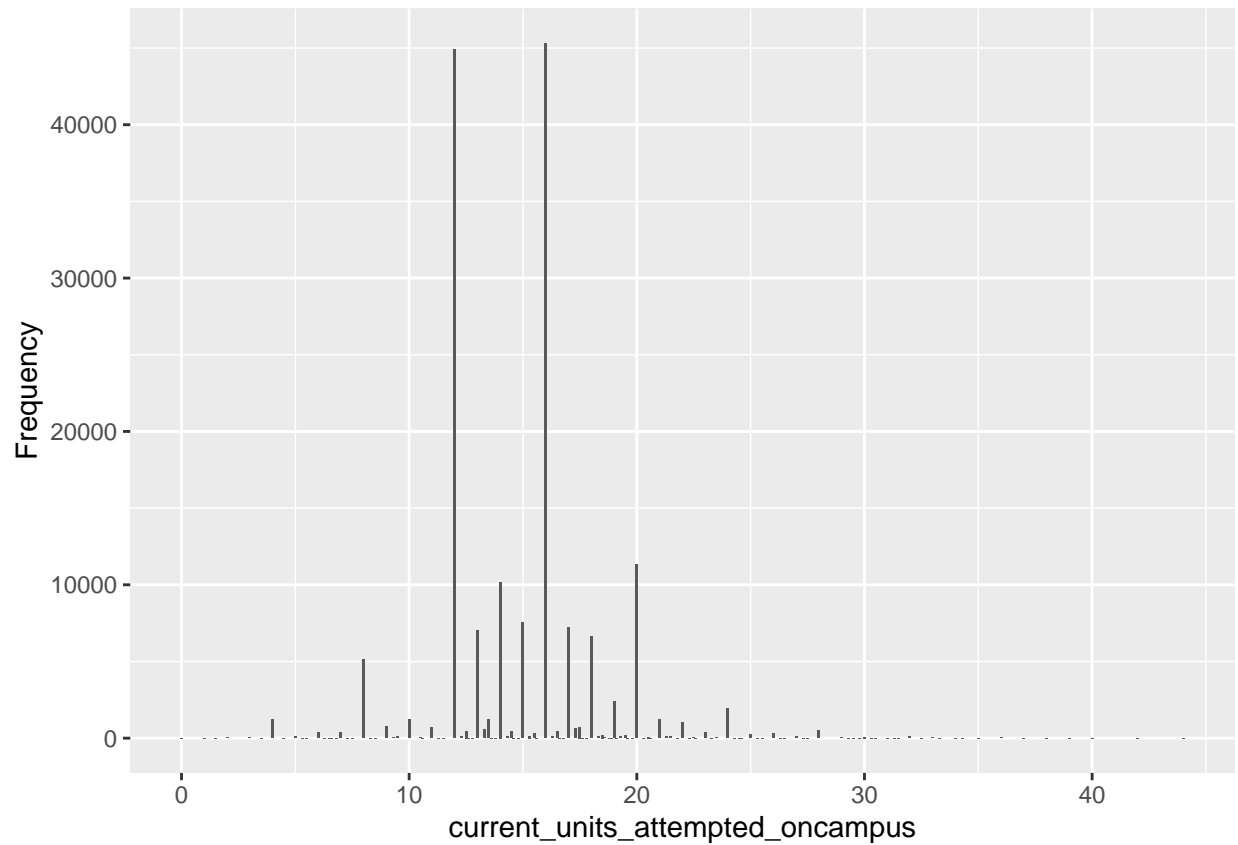
```
[1] -----
[1] Variable: current_units_attempted_online, type: numeric, requested: yes
[1] Values (22 unique): NA, 0, 4, 2, 8, ...
[1] Missing: 81.9%
  Group.1      x
1    2006 0.9990813
2    2007 0.9995796
3    2008 0.9995623
4    2009 0.9992908
5    2010 0.9993285
6    2011 0.9981103
7    2012 0.9972808
8    2013 0.9946479
9    2014 0.9773083
10   2015 0.8593101
11   2016 0.6650633
12   2017 0.4221389
13   2018 0.1244206
```

Warning: Removed 743166 rows containing non-finite values ('stat_count()').



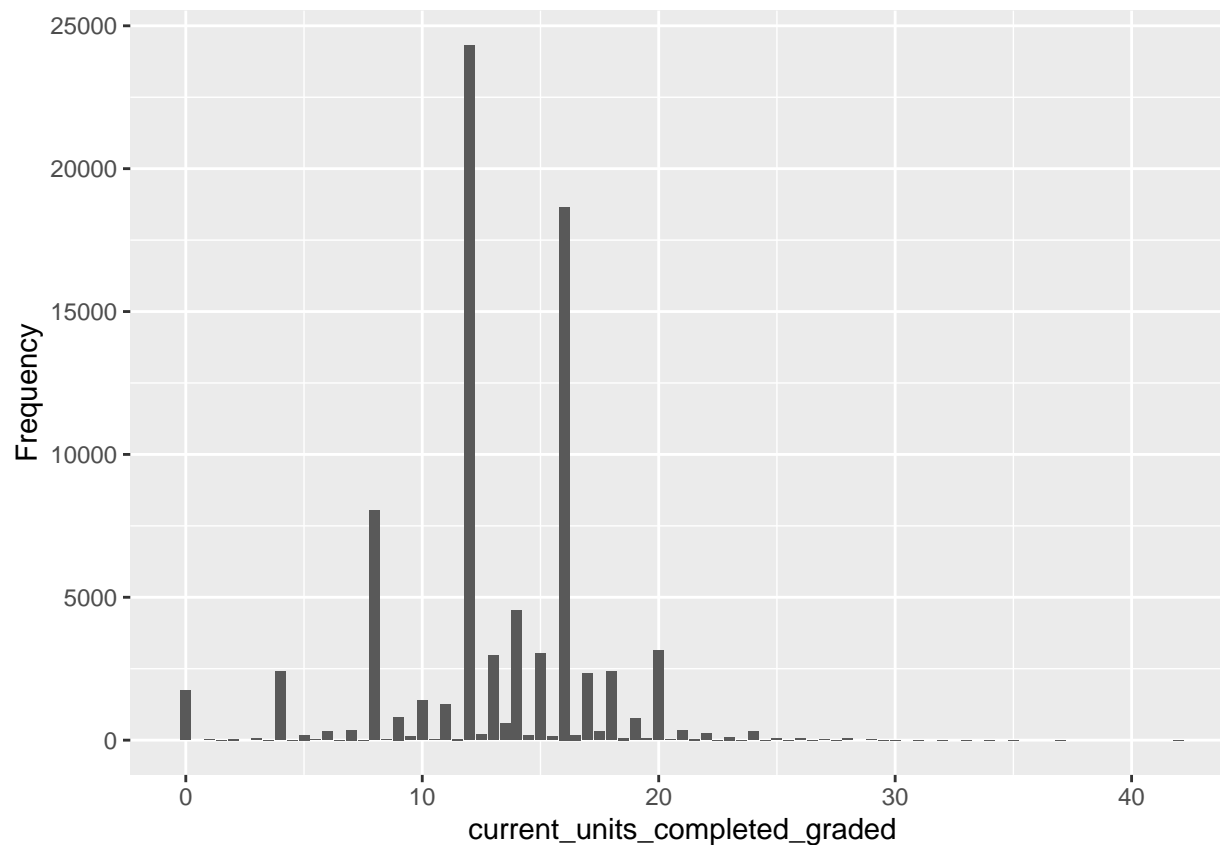
```
[1] -----
[1] Variable: current_units_attempted_oncampus, type: numeric, requested: yes
[1] Values (126 unique): NA, 12, 15, 14, 16, ...
[1] Missing: 81.8%
Group.1      x
1      2006 0.9990107
2      2007 0.9995634
3      2008 0.9995467
4      2009 0.9993081
5      2010 0.9993125
6      2011 0.9980670
7      2012 0.9972808
8      2013 0.9945803
9      2014 0.9771600
10     2015 0.8589683
11     2016 0.6647133
12     2017 0.4210202
13     2018 0.1201100
```

Warning: Removed 742640 rows containing non-finite values ('stat_count()').



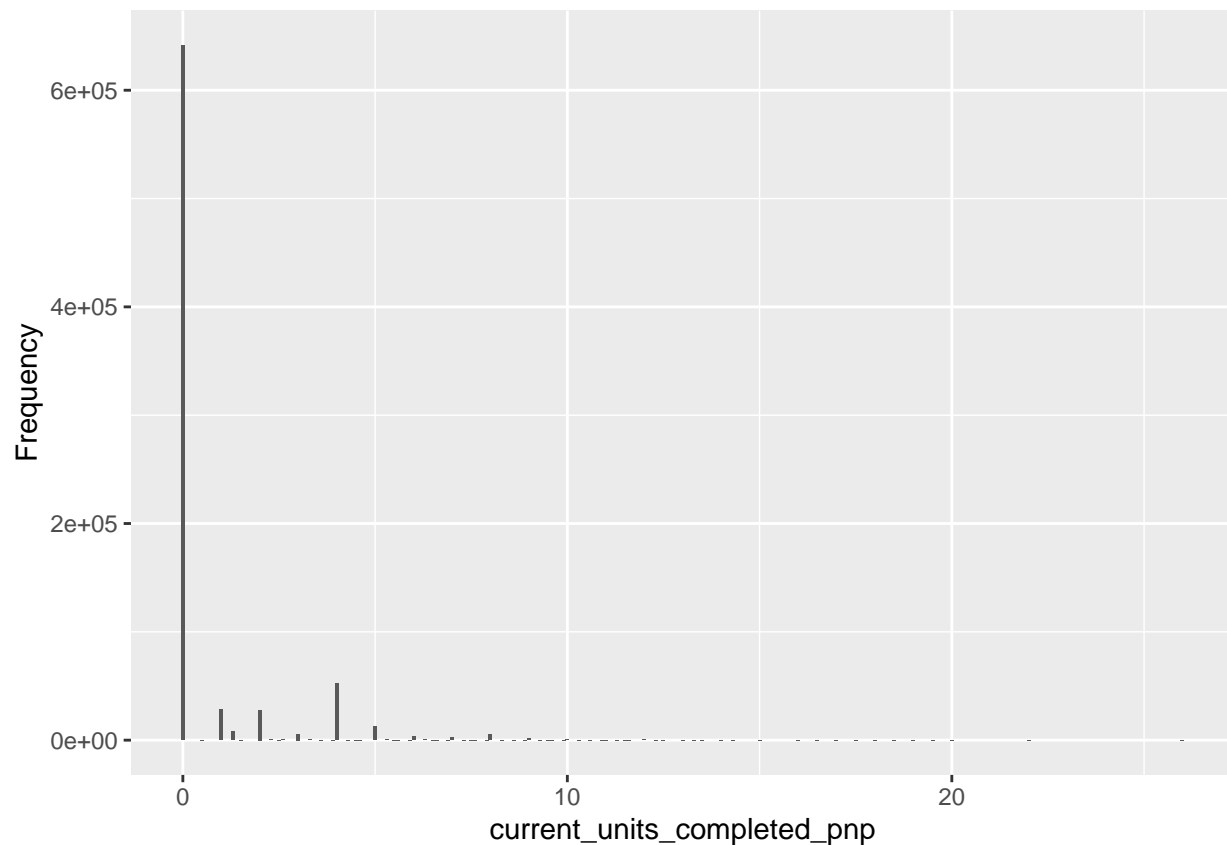
```
[1] -----
[1] Variable: current_units_completed_graded, type: numeric, requested: yes
[1] Values (66 unique): NA, 15, 12, 8, 14, ...
[1] Missing: 91%
[1] Most missing: 2007 100%, Least missing: 2018 55.8%

Warning: Removed 825801 rows containing non-finite values ('stat_count()').
```



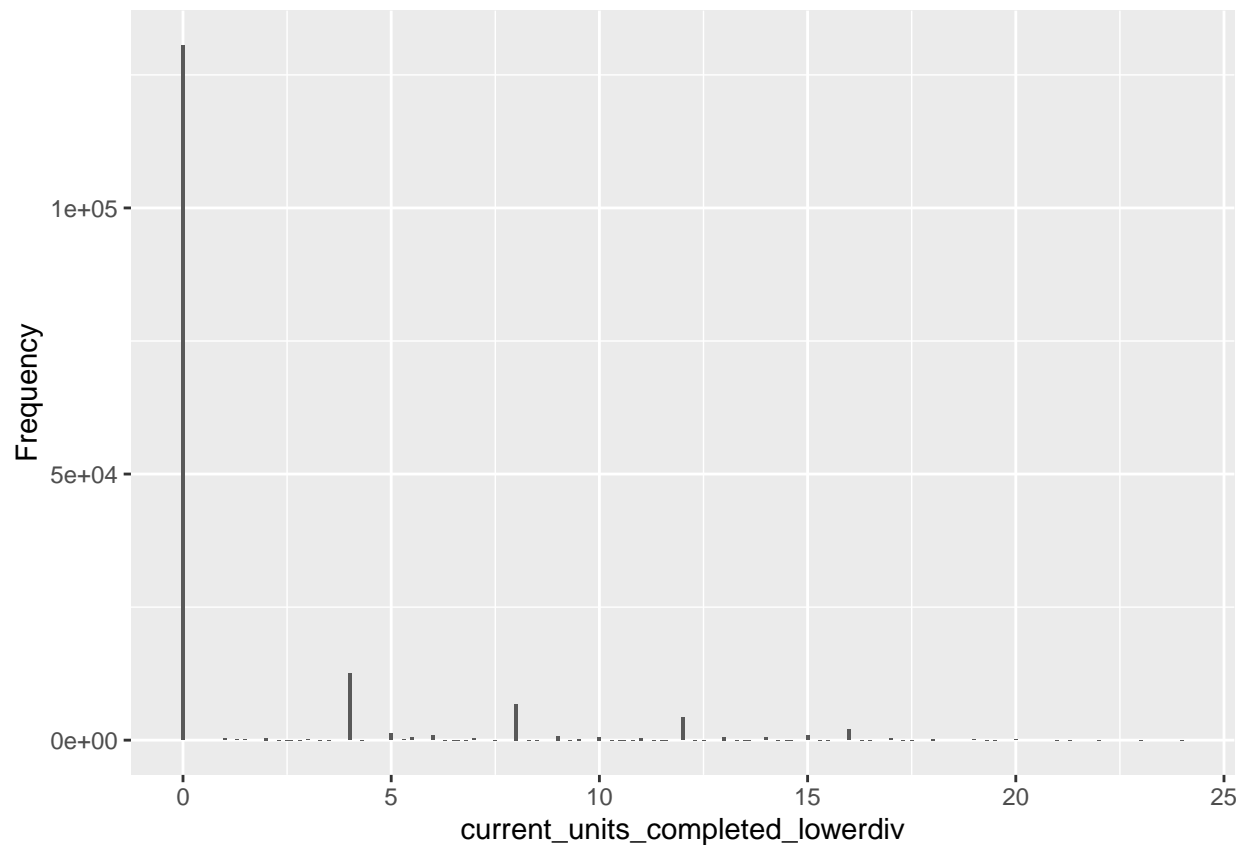
```
[1] -----
[1] Variable: current_units_completed_pnp, type: numeric, requested: yes
[1] Values (70 unique): 0, 6, 1, 1.3, 4, ...
[1] Missing: 12.5%
  Group.1      x
1    2006 0.74395214
2    2007 0.60694707
3    2008 0.39758339
4    2009 0.11914275
5    2010 0.02155179
6    2011 0.02133491
7    2012 0.02067447
8    2013 0.02015137
9    2014 0.01987380
10   2015 0.01797076
11   2016 0.01339114
12   2017 0.01050453
13   2018 0.01924158
```

Warning: Removed 113662 rows containing non-finite values ('stat_count()').



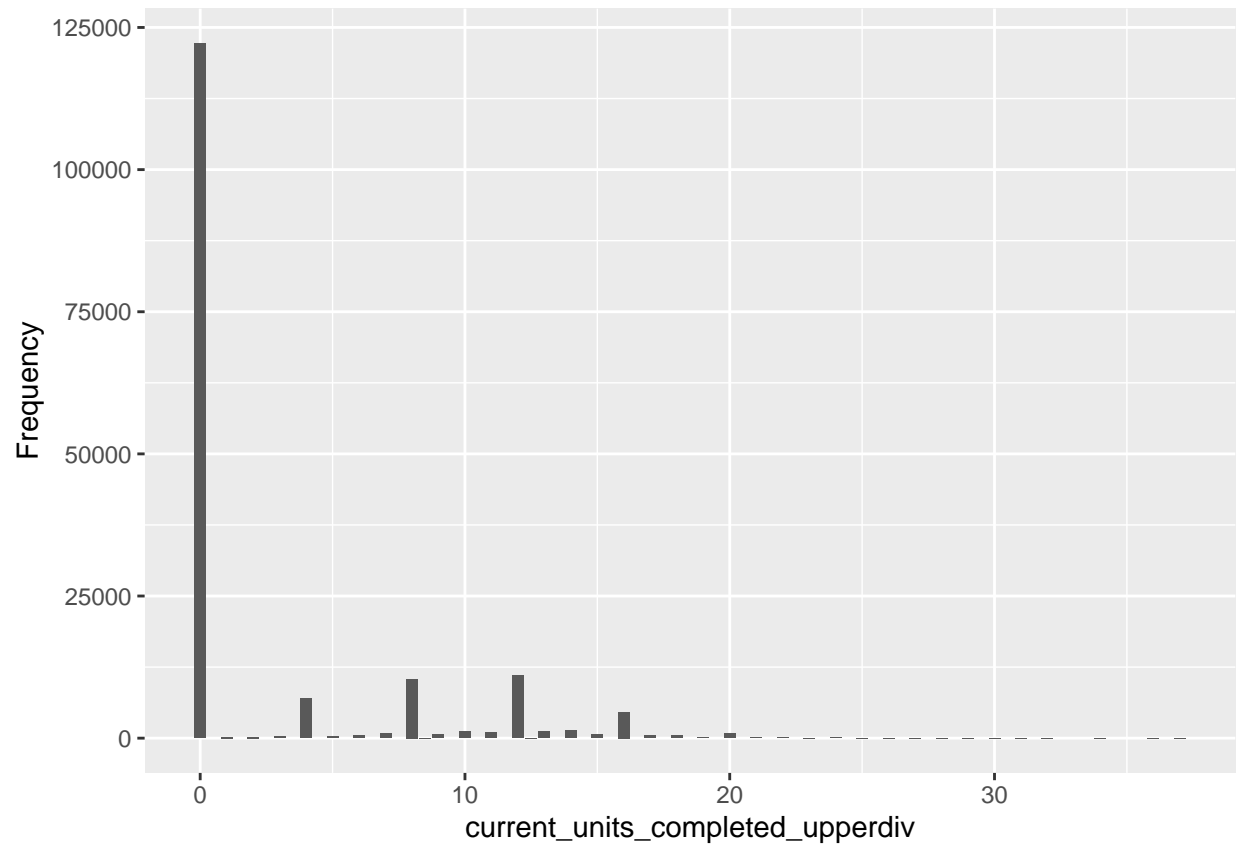
```
[1] -----
[1] Variable: current_units_completed_lowerdiv, type: numeric, requested: yes
[1] Values (70 unique): NA, 8, 0, 4, 5, ...
[1] Missing: 81.8%
  Group.1      x
1    2006 0.9990107
2    2007 0.9995634
3    2008 0.9995623
4    2009 0.9992908
5    2010 0.9993285
6    2011 0.9980959
7    2012 0.9972808
8    2013 0.9945939
9    2014 0.9772004
10   2015 0.8591129
11   2016 0.6647698
12   2017 0.4214342
13   2018 0.1187980
```

Warning: Removed 742595 rows containing non-finite values ('stat_count()').

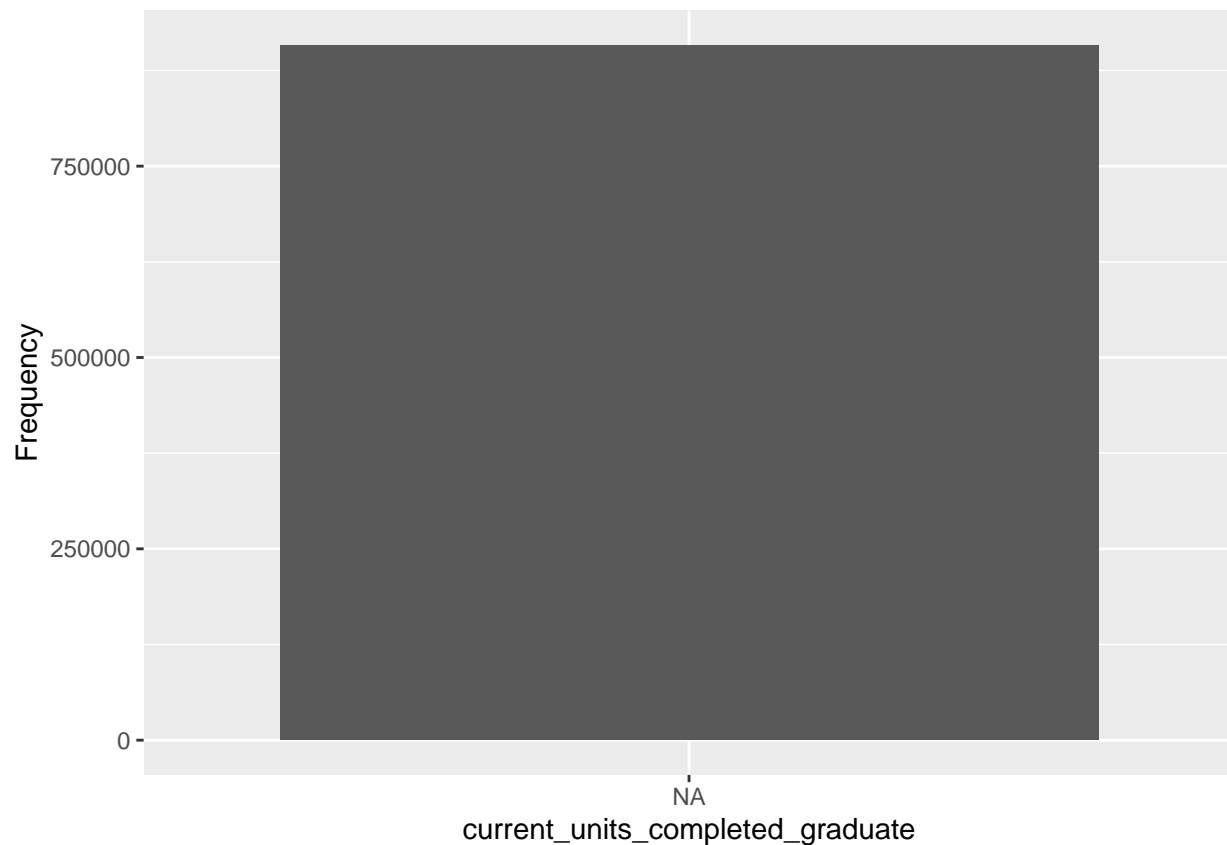


```
[1] -----
[1] Variable: current_units_completed_upperdiv, type: numeric, requested: yes
[1] Values (39 unique): NA, 4, 15, 0, 14, ...
[1] Missing: 81.8%
  Group.1      x
1    2006 0.9990107
2    2007 0.9995634
3    2008 0.9995467
4    2009 0.9992735
5    2010 0.9993125
6    2011 0.9980670
7    2012 0.9972808
8    2013 0.9945939
9    2014 0.9772004
10   2015 0.8588894
11   2016 0.6646456
12   2017 0.4210986
13   2018 0.1160492
```

Warning: Removed 742312 rows containing non-finite values ('stat_count()').

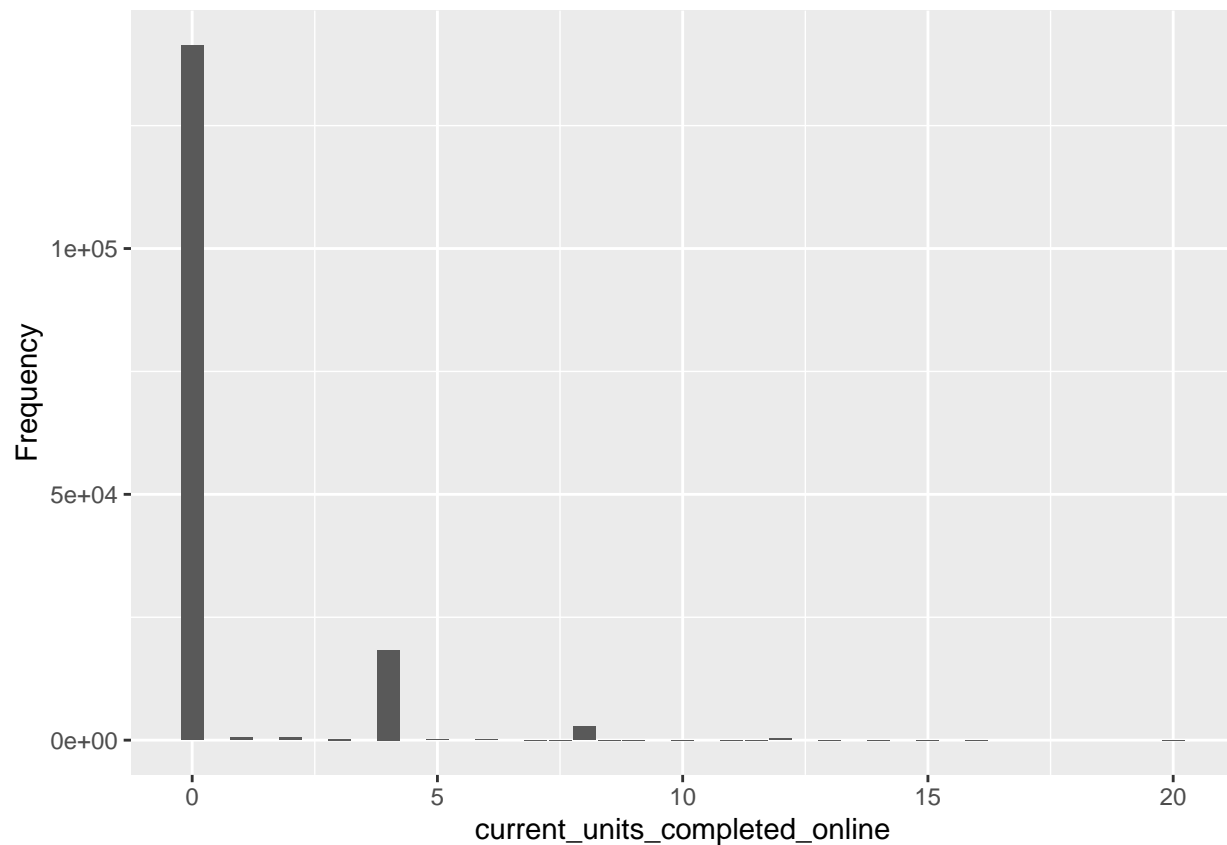


```
[1] -----  
[1] Variable: current_units_completed_graduate, type: logical, requested: yes  
[1] Values (1 unique): NA  
[1] Missing: 100%
```



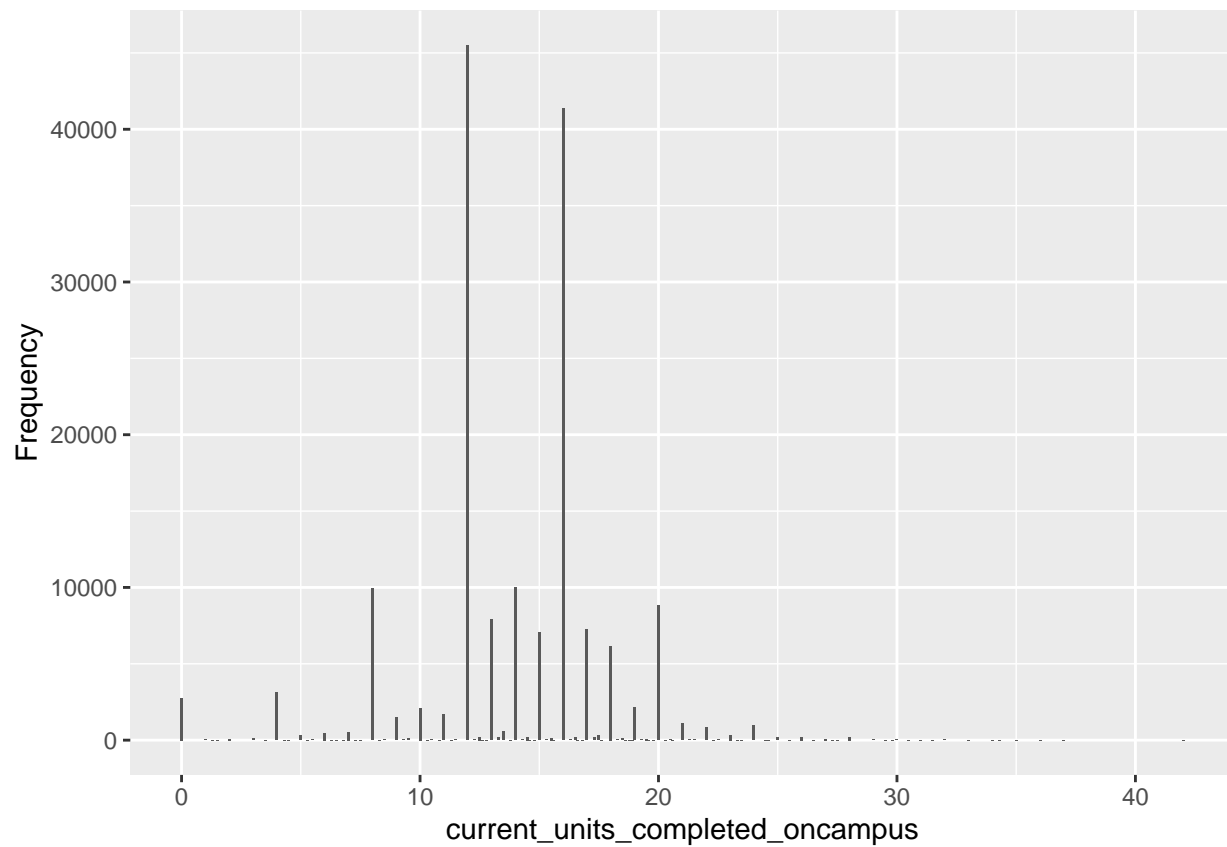
```
[1] -----
[1] Variable: current_units_completed_online, type: numeric, requested: yes
[1] Values (22 unique): NA, 0, 4, 2, 8, ...
[1] Missing: 81.9%
  Group.1      x
1    2006 0.9990813
2    2007 0.9995796
3    2008 0.9995623
4    2009 0.9992908
5    2010 0.9993285
6    2011 0.9981103
7    2012 0.9972808
8    2013 0.9946479
9    2014 0.9773083
10   2015 0.8593101
11   2016 0.6650633
12   2017 0.4221389
13   2018 0.1244206
```

Warning: Removed 743166 rows containing non-finite values ('stat_count()').

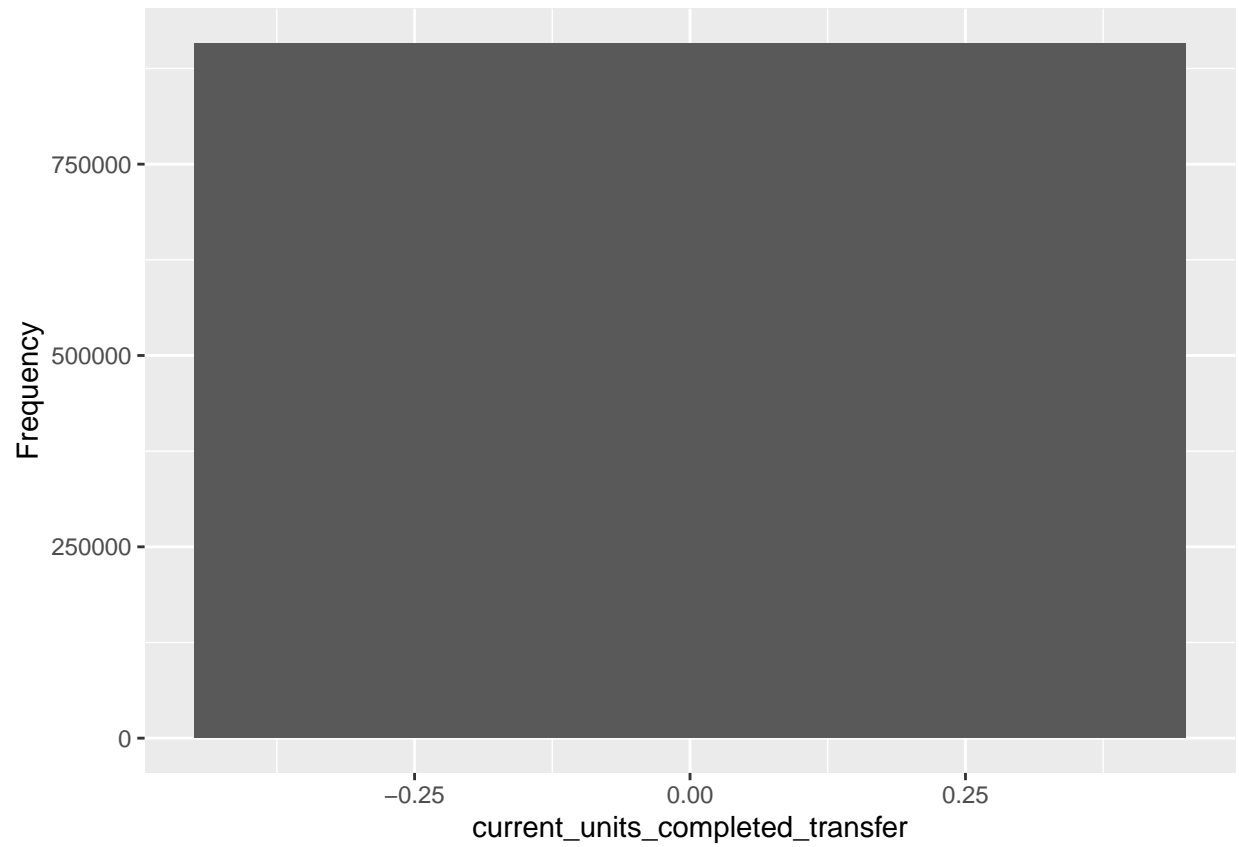


```
[1] -----
[1] Variable: current_units_completed_oncampus, type: numeric, requested: yes
[1] Values (111 unique): NA, 12, 15, 14, 16, ...
[1] Missing: 81.8%
  Group.1      x
1    2006 0.9990107
2    2007 0.9995634
3    2008 0.9995467
4    2009 0.9992735
5    2010 0.9993125
6    2011 0.9980670
7    2012 0.9972808
8    2013 0.9945803
9    2014 0.9771734
10   2015 0.8588894
11   2016 0.6646569
12   2017 0.4209979
13   2018 0.1157369
```

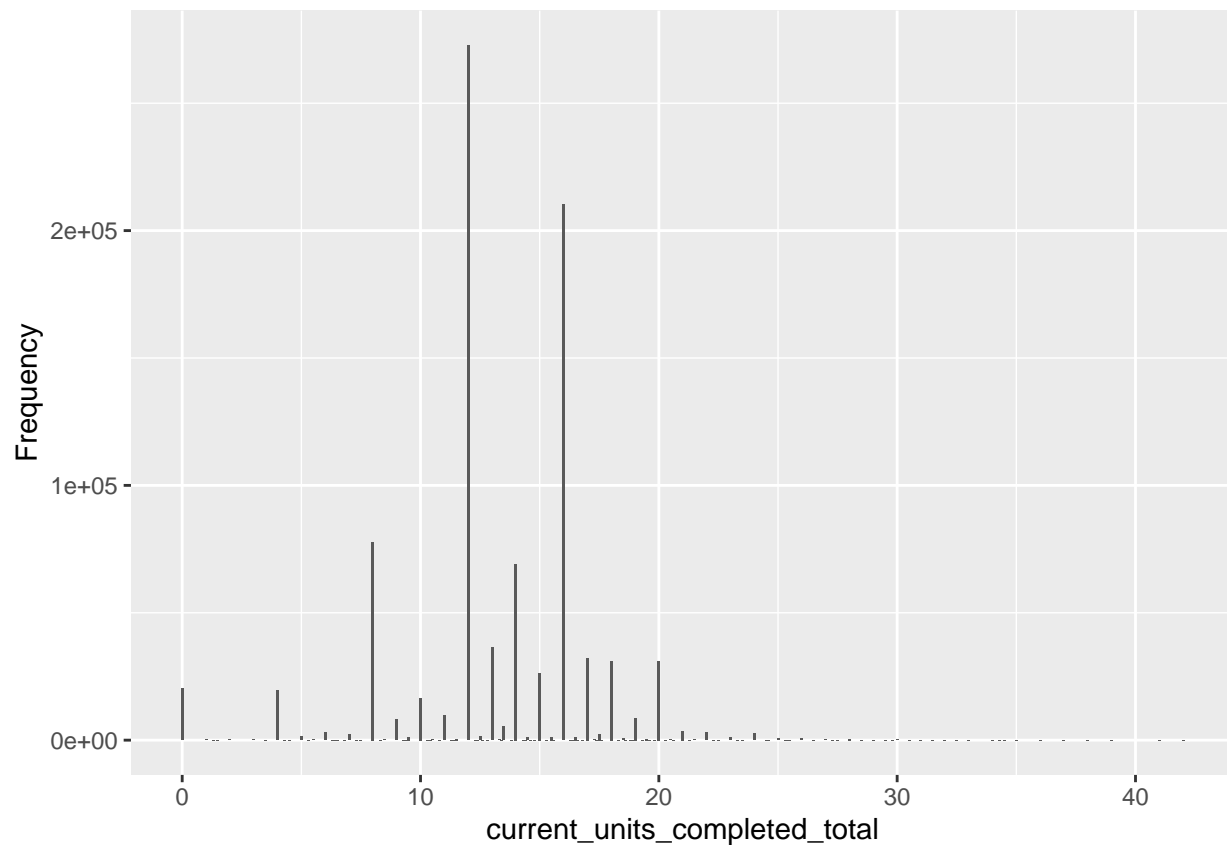
Warning: Removed 742276 rows containing non-finite values ('stat_count()').



```
[1] -----  
[1] Variable: current_units_completed_transfer, type: numeric, requested: yes  
[1] Values (1 unique): 0  
[1] Missing: 0%
```

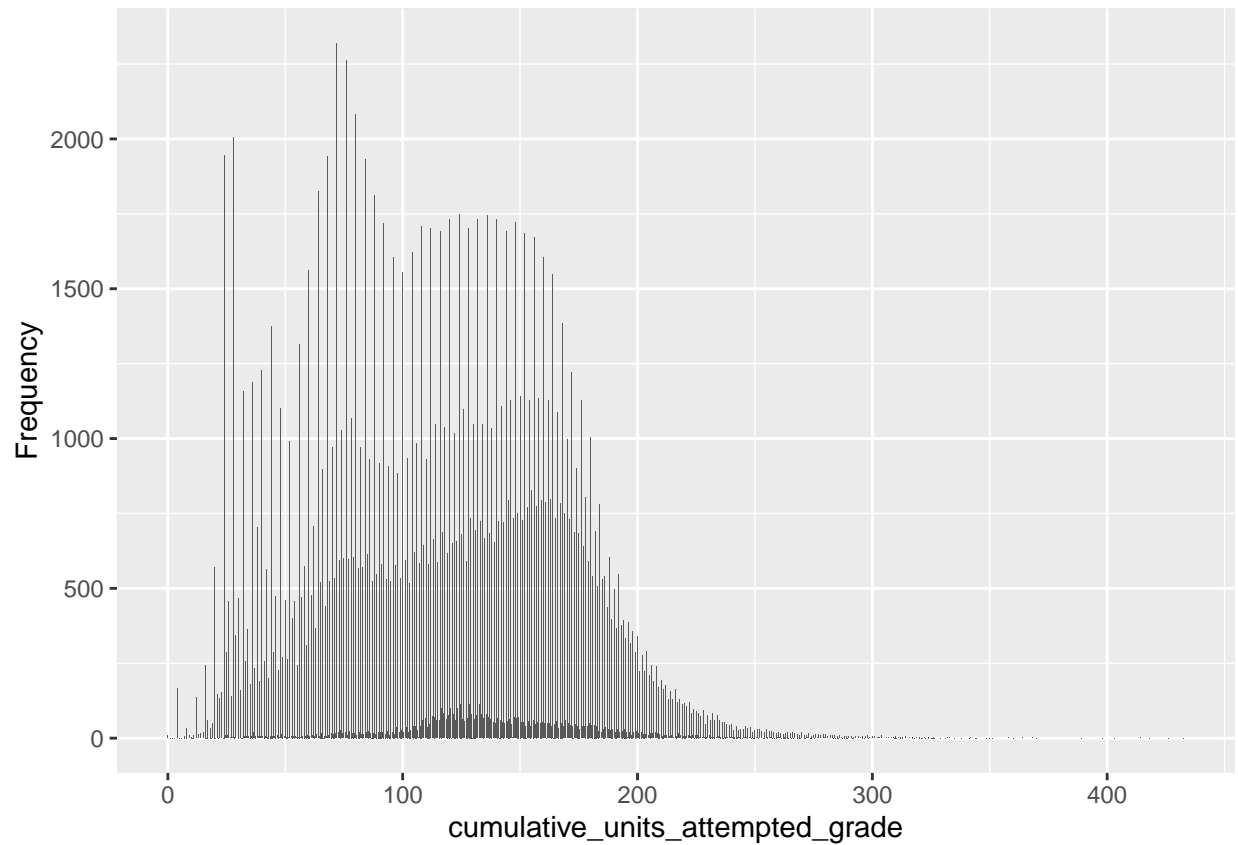


```
[1] -----  
[1] Variable: current_units_completed_total, type: numeric, requested: yes  
[1] Values (127 unique): 12, 15, 14, 17, 13, ...  
[1] Missing: 0%
```



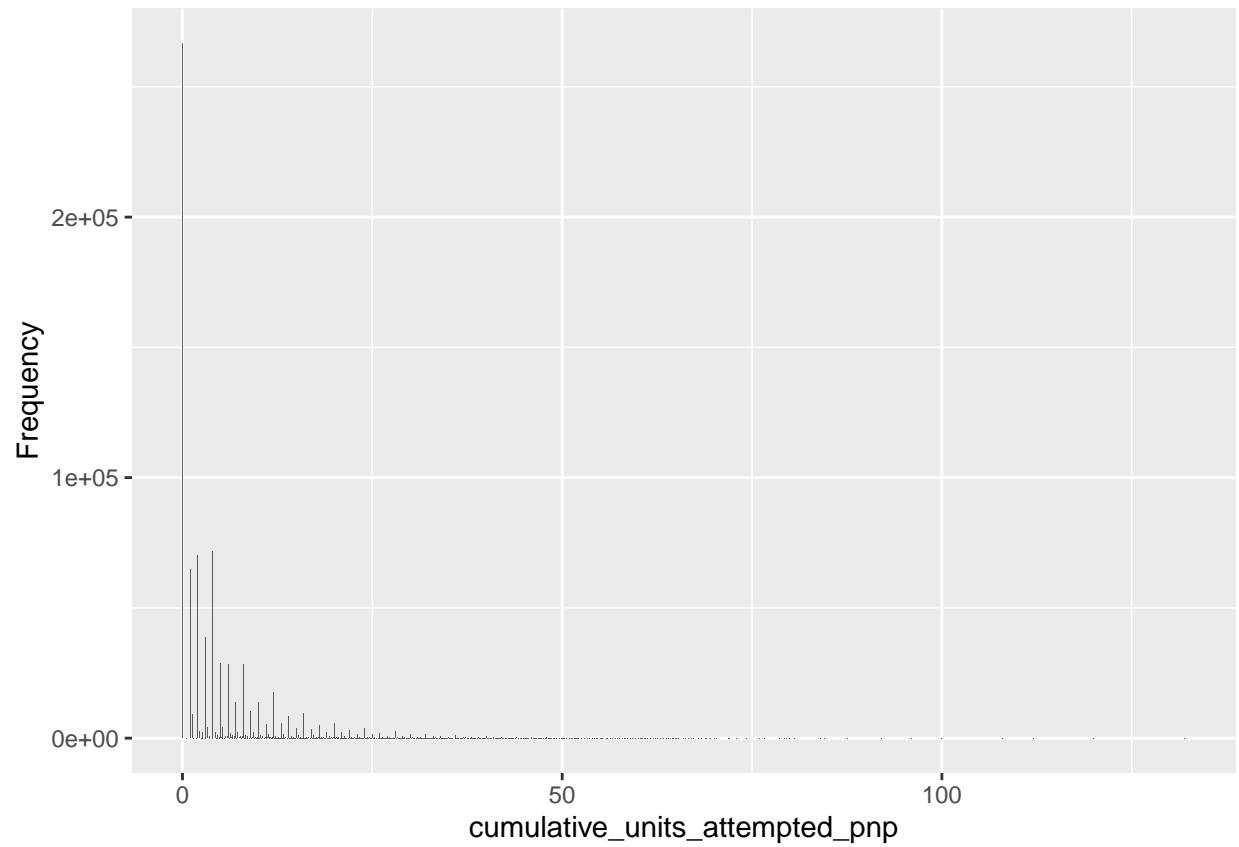
```
[1] -----
[1] Variable: cumulative_units_attempted_grade, type: numeric, requested: yes
[1] Values (712 unique): NA, 64, 79, 36, 151, ...
[1] Missing: 81.7%
  Group.1      x
1    2006 0.9989164
2    2007 0.9995310
3    2008 0.9995311
4    2009 0.9992562
5    2010 0.9993125
6    2011 0.9980093
7    2012 0.9972217
8    2013 0.9945128
9    2014 0.9770521
10   2015 0.8589157
11   2016 0.6668134
12   2017 0.4215572
13   2018 0.1076654
```

Warning: Removed 741843 rows containing non-finite values ('stat_count()').

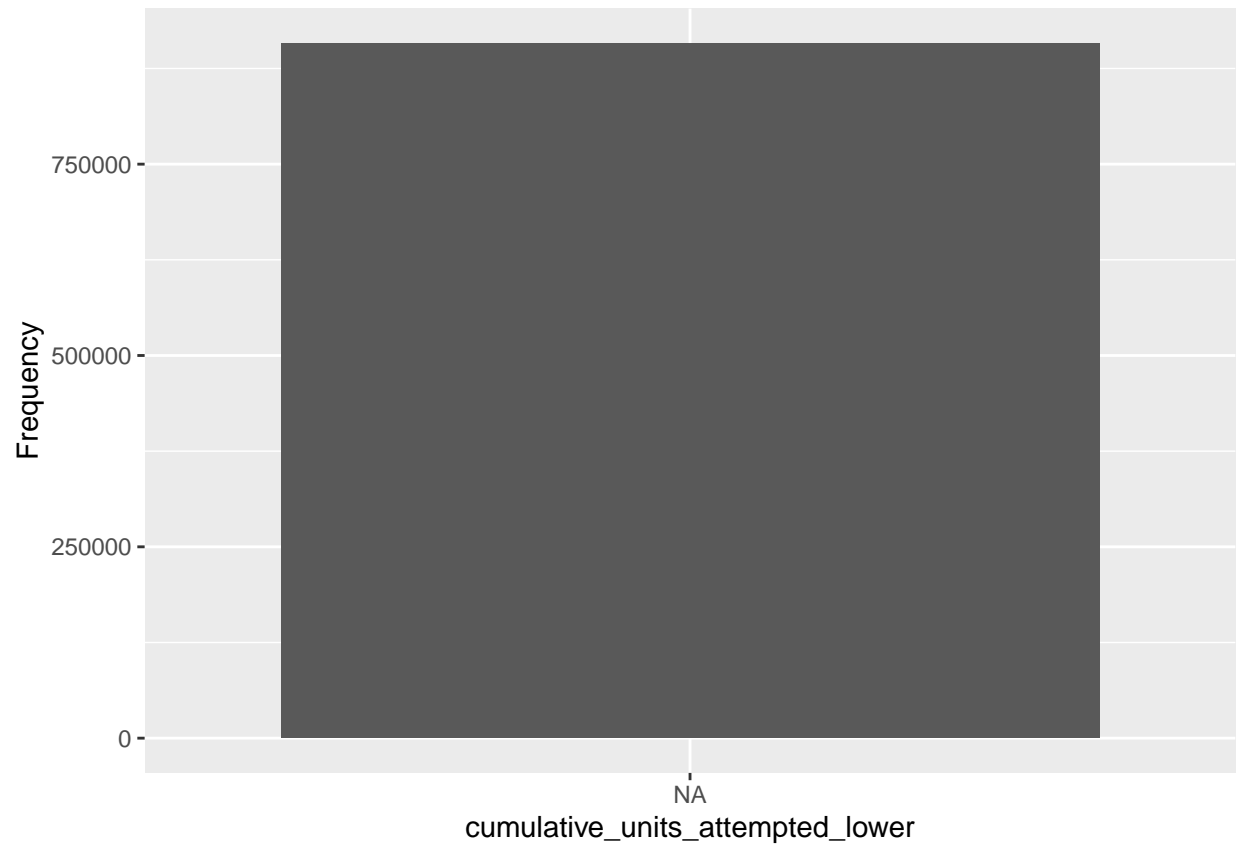


```
[1] -----
[1] Variable: cumulative_units_attempted_pnp, type: numeric, requested: yes
[1] Values (484 unique): 0, 6, 1.3, 2.3, 8, ...
[1] Missing: 12.3%
  Group.1      x
1    2006 0.743810803
2    2007 0.606914730
3    2008 0.397552131
4    2009 0.119125456
5    2010 0.021535805
6    2011 0.021233934
7    2012 0.020615357
8    2013 0.020002703
9    2014 0.019496279
10   2015 0.017037388
11   2016 0.012397534
12   2017 0.007607115
13   2018 0.001349410
```

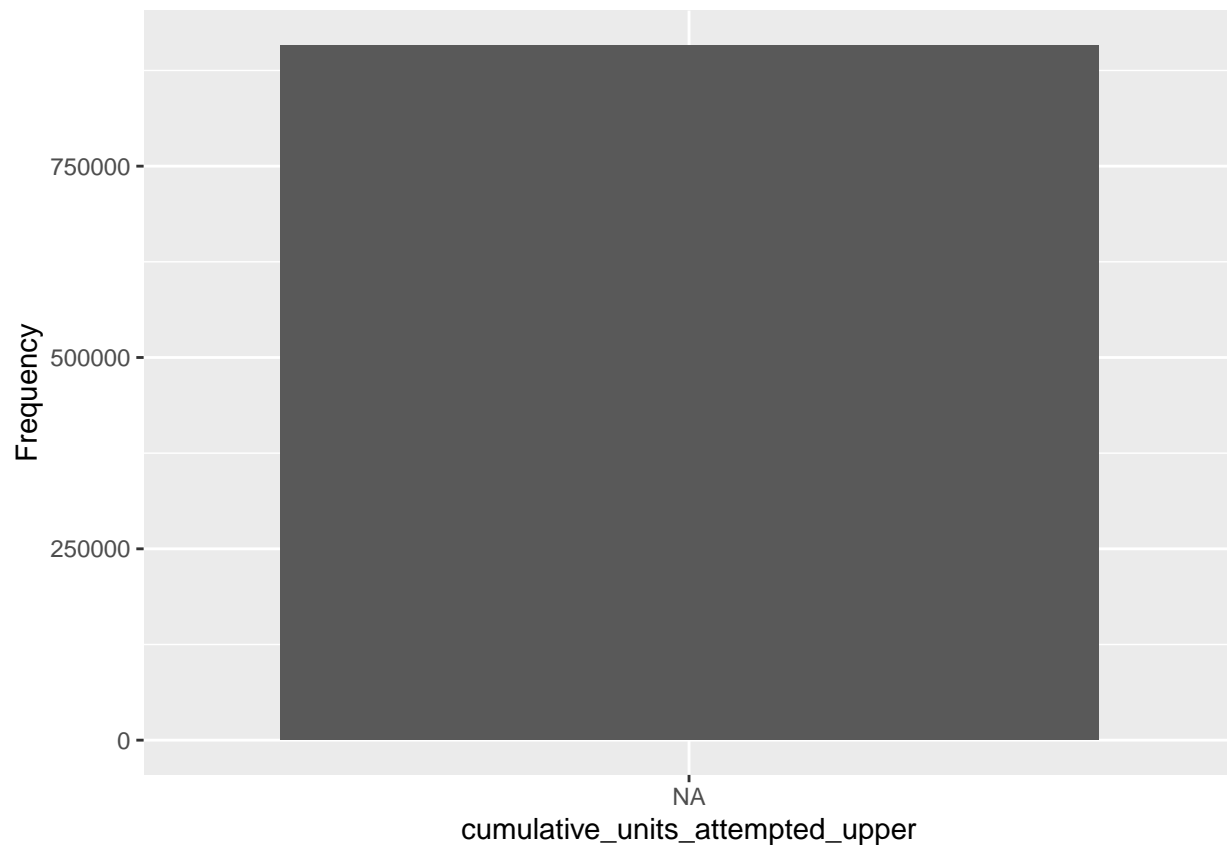
Warning: Removed 111750 rows containing non-finite values ('stat_count()').



```
[1] -----  
[1] Variable: cumulative_units_attempted_lower, type: logical, requested: yes  
[1] Values (1 unique): NA  
[1] Missing: 100%
```

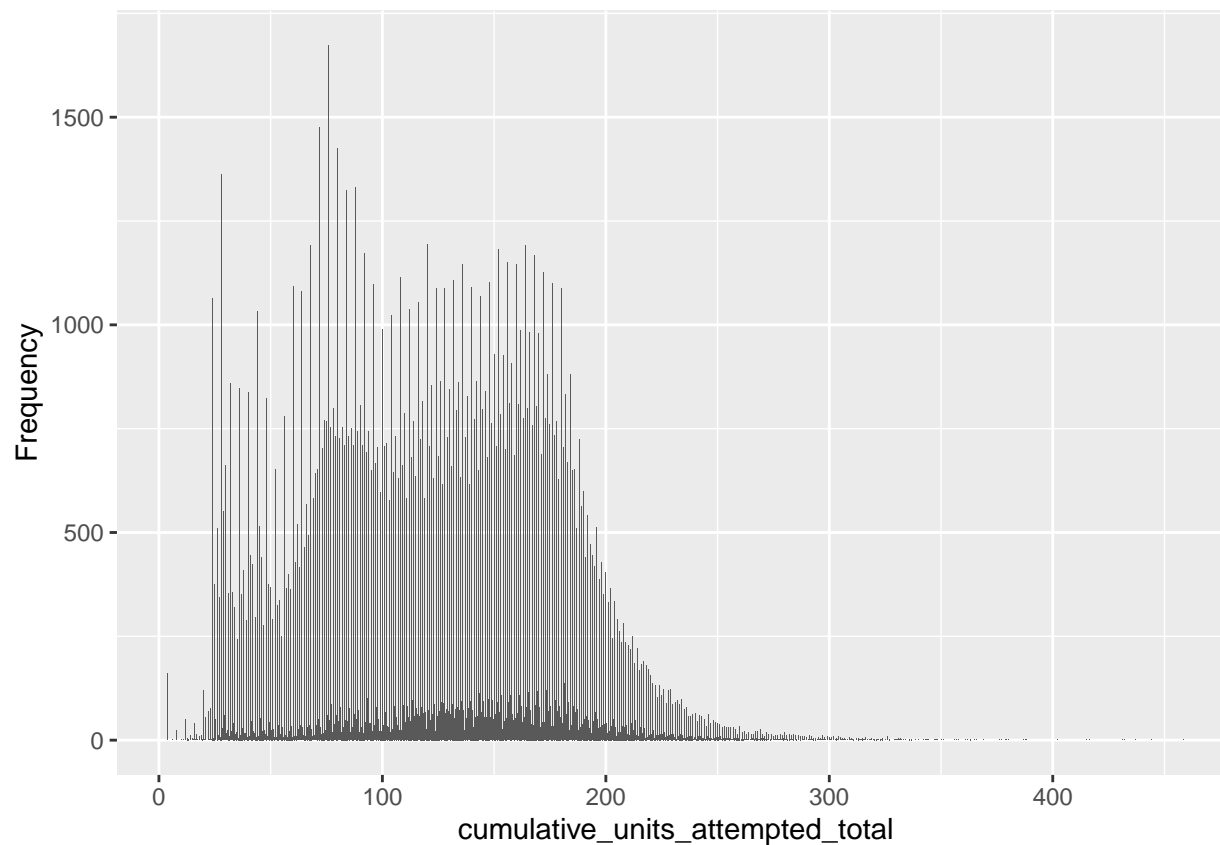


```
[1] -----  
[1] Variable: cumulative_units_attempted_upper, type: logical, requested: yes  
[1] Values (1 unique): NA  
[1] Missing: 100%
```



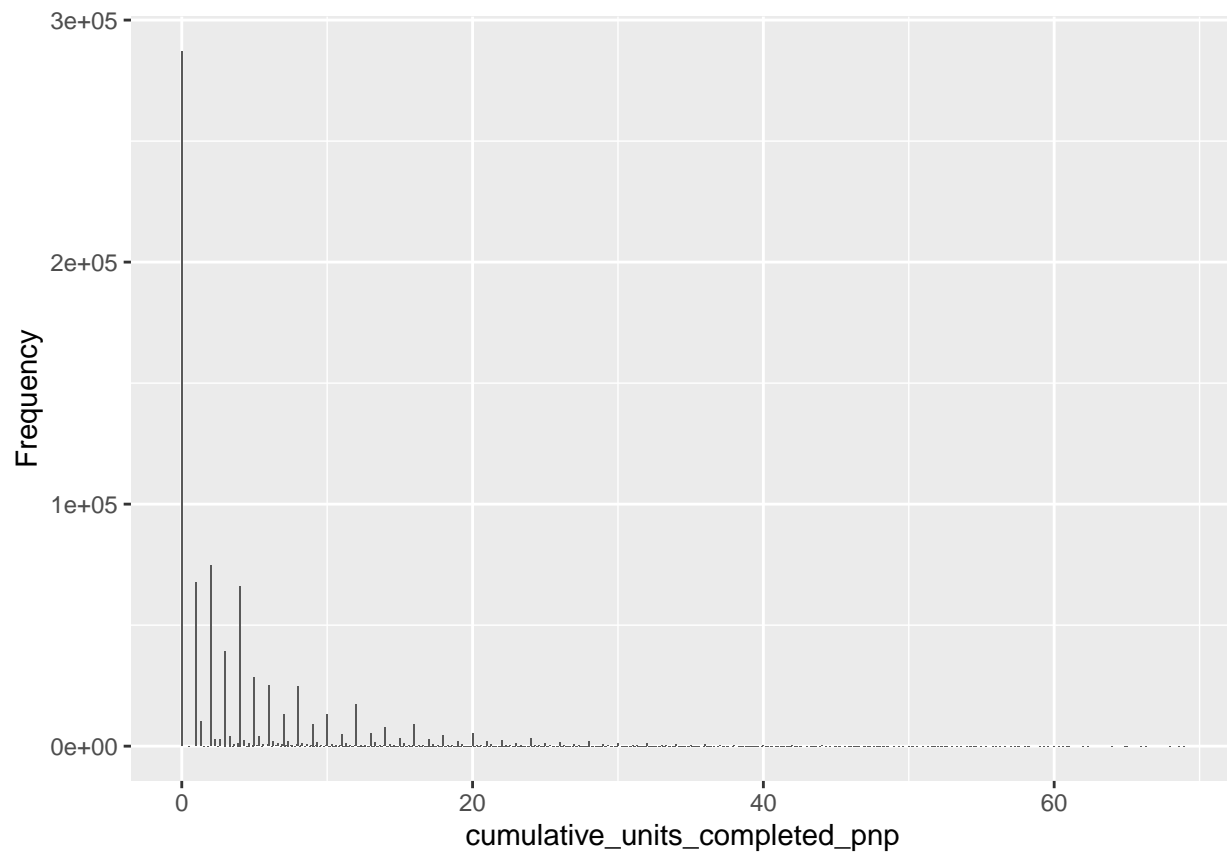
```
[1] -----
[1] Variable: cumulative_units_attempted_total, type: numeric, requested: yes
[1] Values (2070 unique): NA, 64, 79, 36, 151, ...
[1] Missing: 81.7%
  Group.1      x
1    2006 0.9989164
2    2007 0.9995310
3    2008 0.9995467
4    2009 0.9992562
5    2010 0.9993125
6    2011 0.9980237
7    2012 0.9972365
8    2013 0.9945398
9    2014 0.9770791
10   2015 0.8589551
11   2016 0.6670393
12   2017 0.4217698
13   2018 0.1081152
```

Warning: Removed 741928 rows containing non-finite values ('stat_count()').

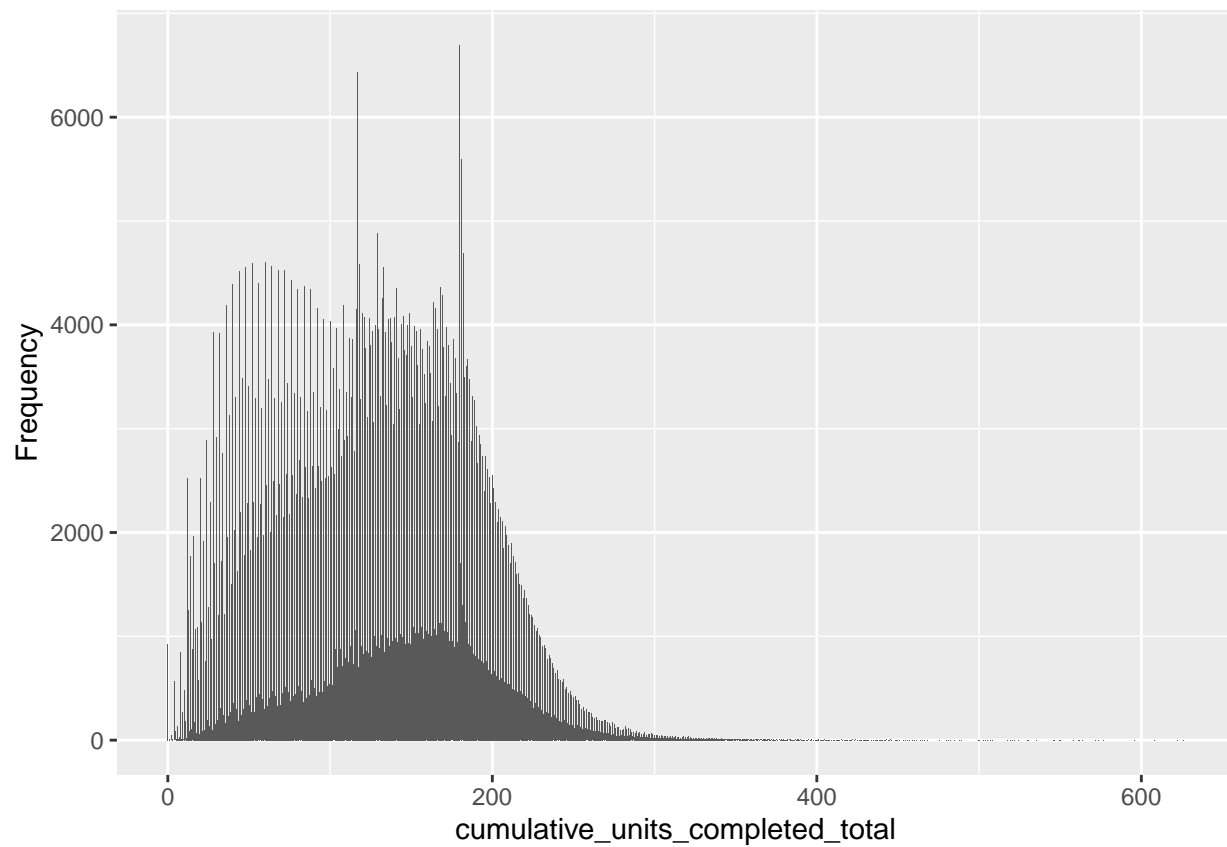


```
[1] -----
[1] Variable: cumulative_units_completed_pnp, type: numeric, requested: yes
[1] Values (431 unique): 0, 6, 1.3, 2.3, 8, ...
[1] Missing: 12.3%
  Group.1      x
1    2006 0.743810803
2    2007 0.606914730
3    2008 0.397552131
4    2009 0.119125456
5    2010 0.021535805
6    2011 0.021233934
7    2012 0.020615357
8    2013 0.020002703
9    2014 0.019496279
10   2015 0.017037388
11   2016 0.012397534
12   2017 0.007607115
13   2018 0.001349410
```

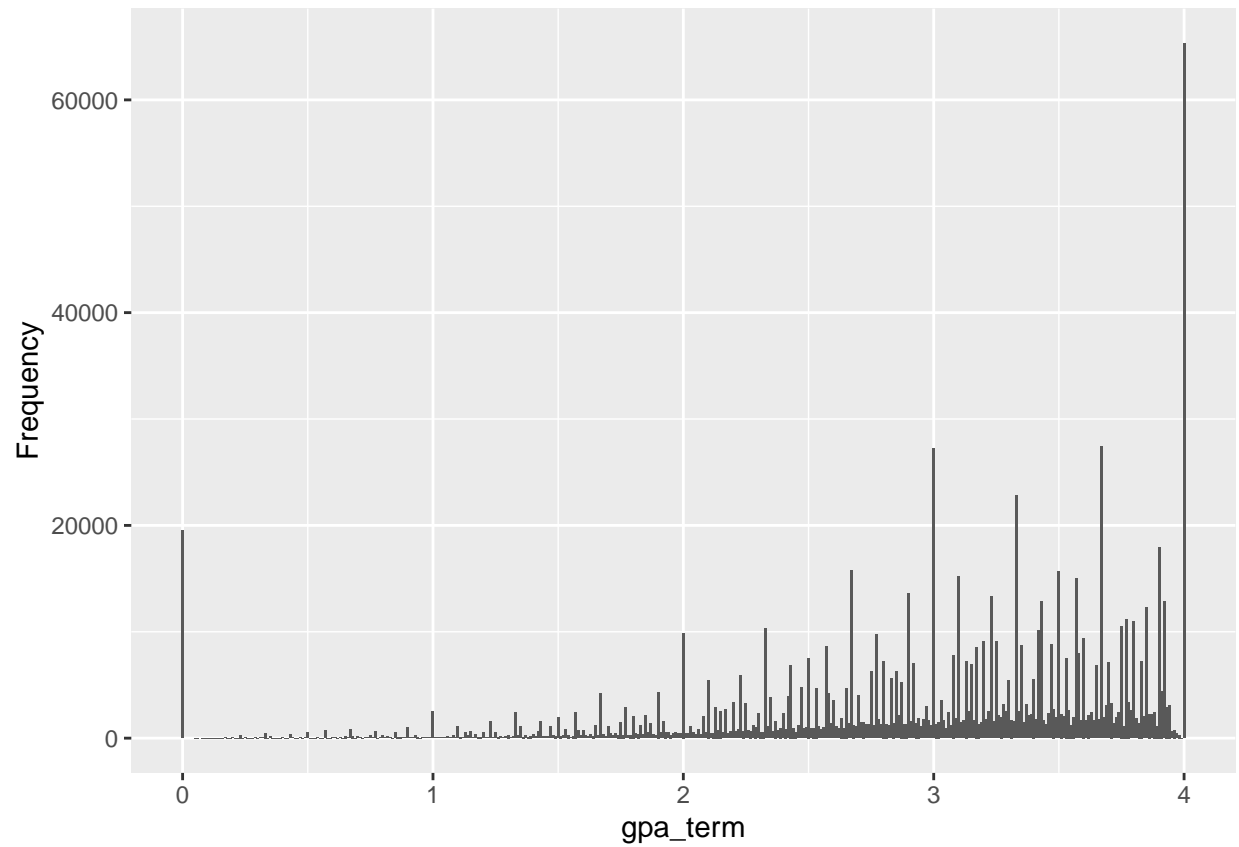
Warning: Removed 111750 rows containing non-finite values ('stat_count()').



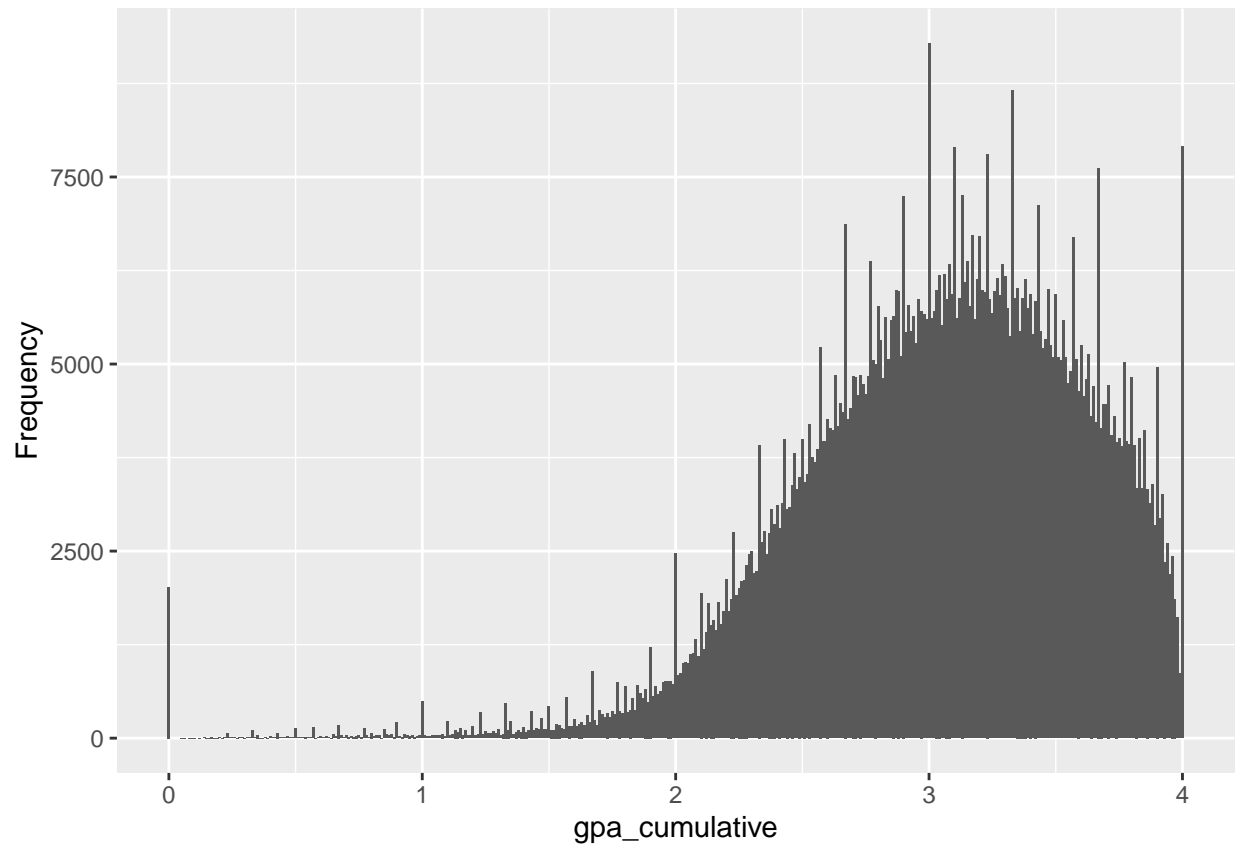
```
[1] -----  
[1] Variable: cumulative_units_completed_total, type: numeric, requested: yes  
[1] Values (3411 unique): 44, 96, 111, 68, 183, ...  
[1] Missing: 0%
```



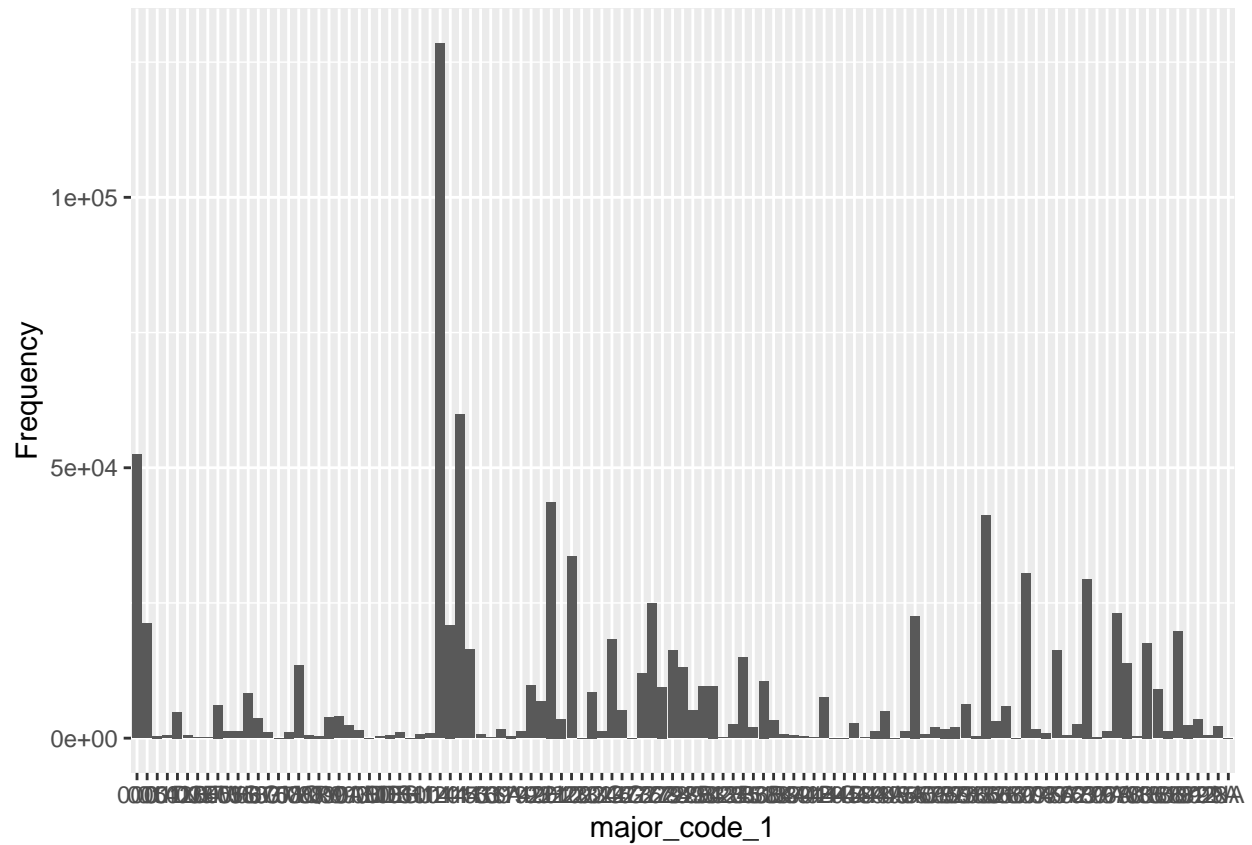
```
[1] -----  
[1] Variable: gpa_term, type: numeric, requested: yes  
[1] Values (396 unique): 3.4300001, 3.53, 3.6500001, 3.23, 4, ...  
[1] Missing: 0%
```



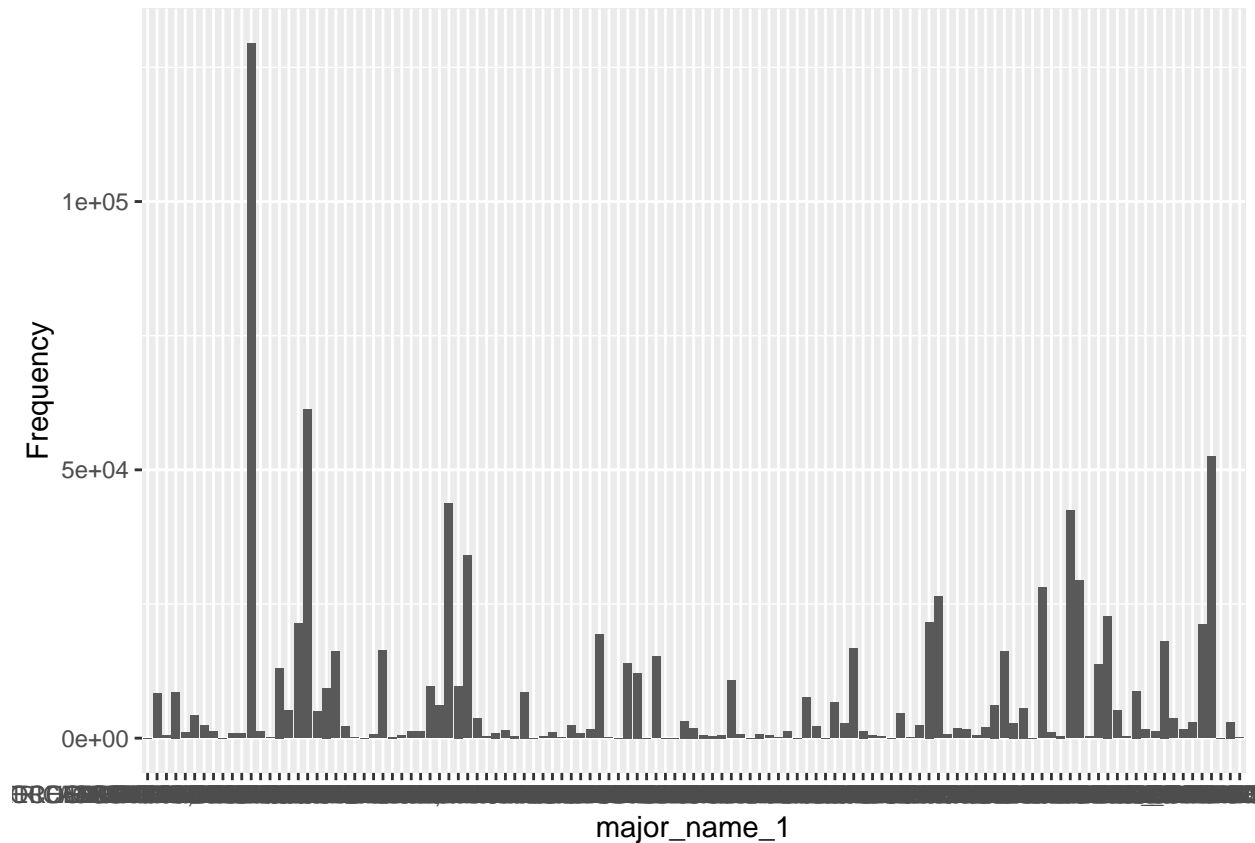
```
[1] -----  
[1] Variable: gpa_cumulative, type: numeric, requested: yes  
[1] Values (394 unique): 3.4300001, 3.3499999, 3.4100001, 3.26, 3.5799999, ...  
[1] Missing: 0%
```



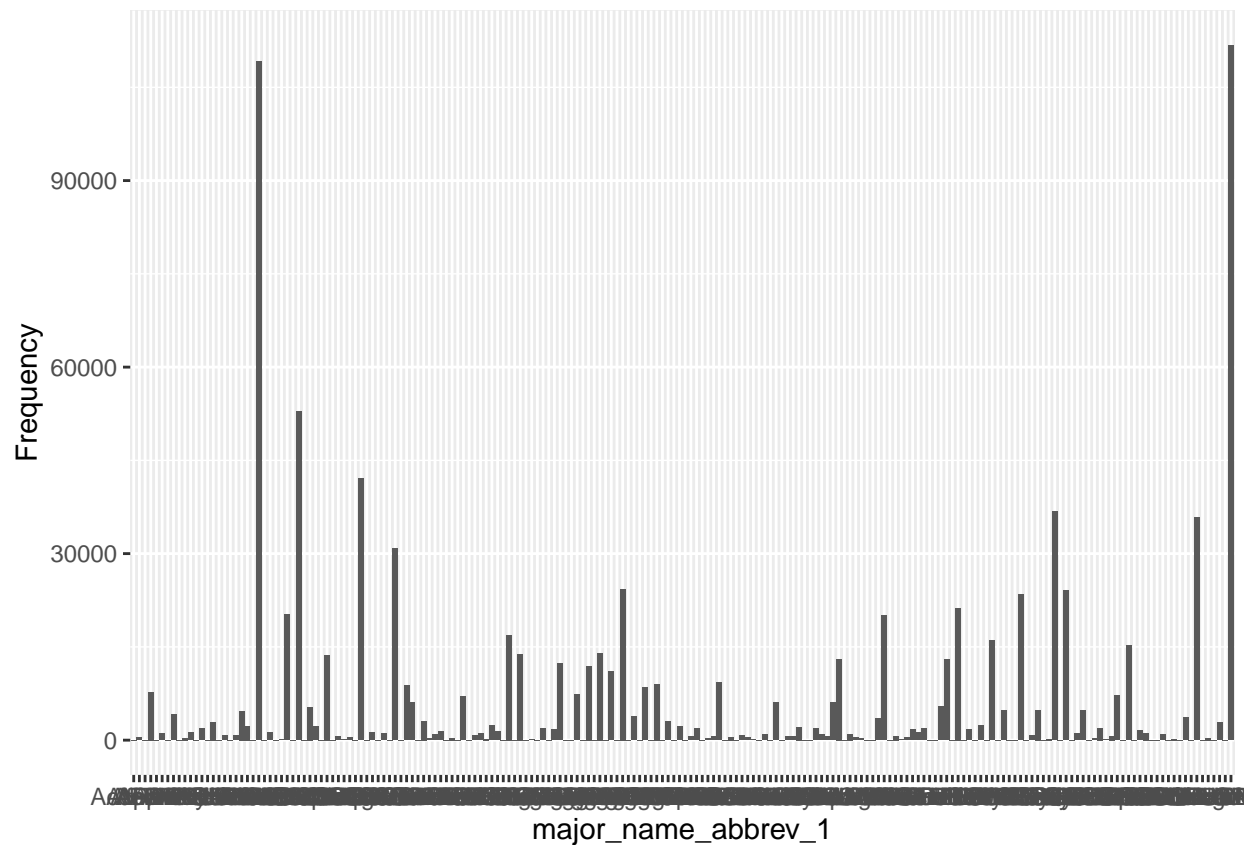
```
[1] -----  
[1] Variable: major_code_1, type: character, requested: yes  
[1] Values (109 unique): 598, 201, 005, 000, 867, ...  
[1] Missing: 0%
```

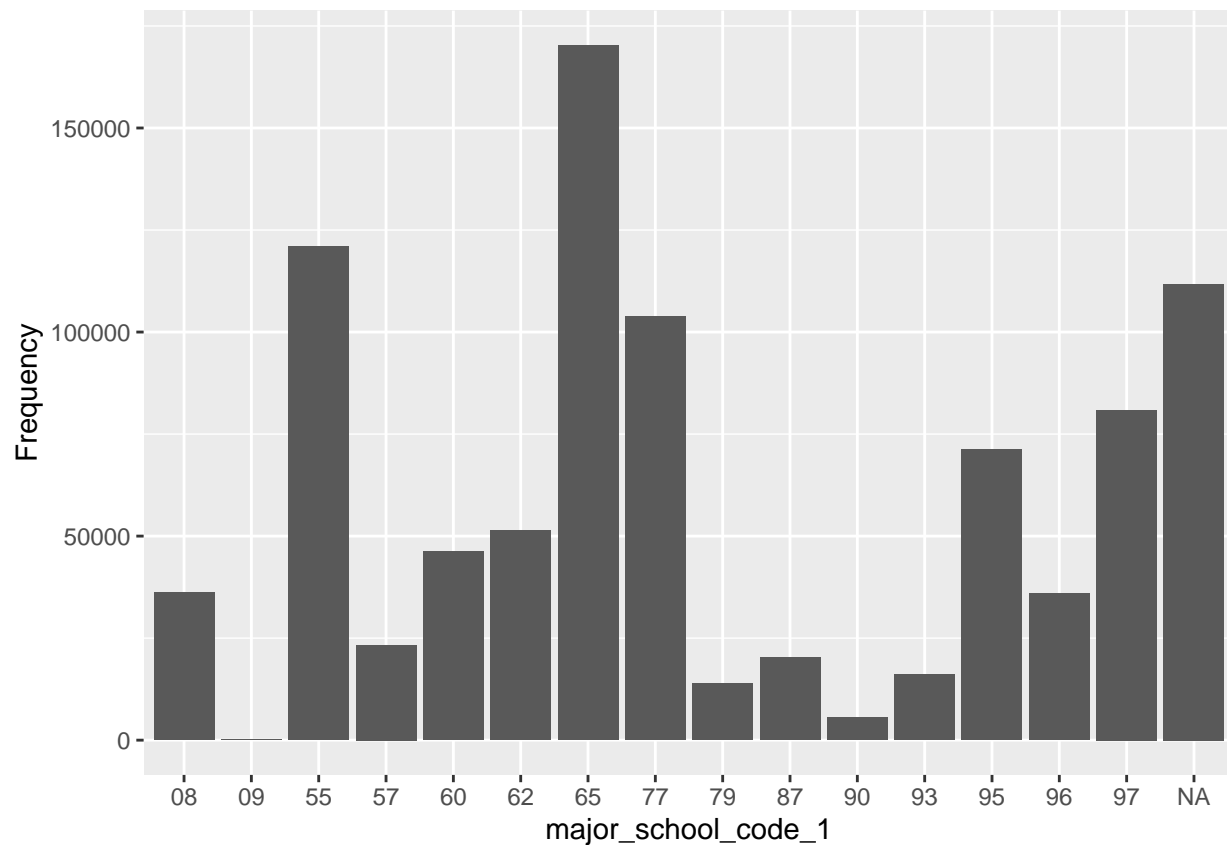
```
[1] -----
[1] Variable: major_name_1, type: character, requested: yes
[1] Values (117 unique): NURSING SCIENCE, COMPUTER SCIENCE, UNAFFILIATED, UNDECLARED, SOCIOLOGY, ...
[1] Missing: 0%
```



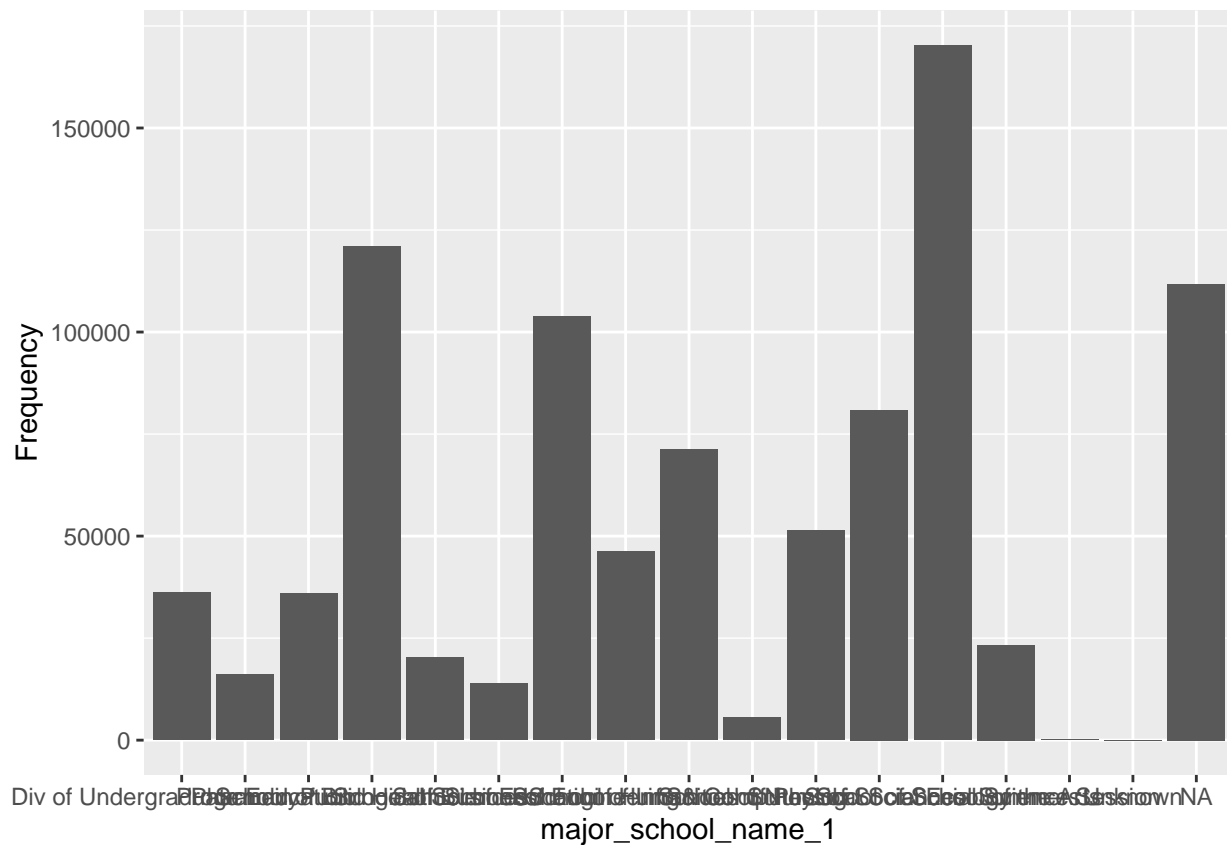
```
[1] -----
[1] Variable: major_name_abbrev_1, type: character, requested: yes
[1] Values (194 unique): Nur Sci, CmptSci, ICS-Ua, Una-Ud, Sociol, ...
[1] Missing: 12.3%
  Group.1      x
1    2006 0.743810803
2    2007 0.606914730
3    2008 0.397552131
4    2009 0.119125456
5    2010 0.021535805
6    2011 0.021233934
7    2012 0.020615357
8    2013 0.020002703
9    2014 0.019496279
10   2015 0.017037388
11   2016 0.012397534
12   2017 0.007607115
13   2018 0.001349410
```



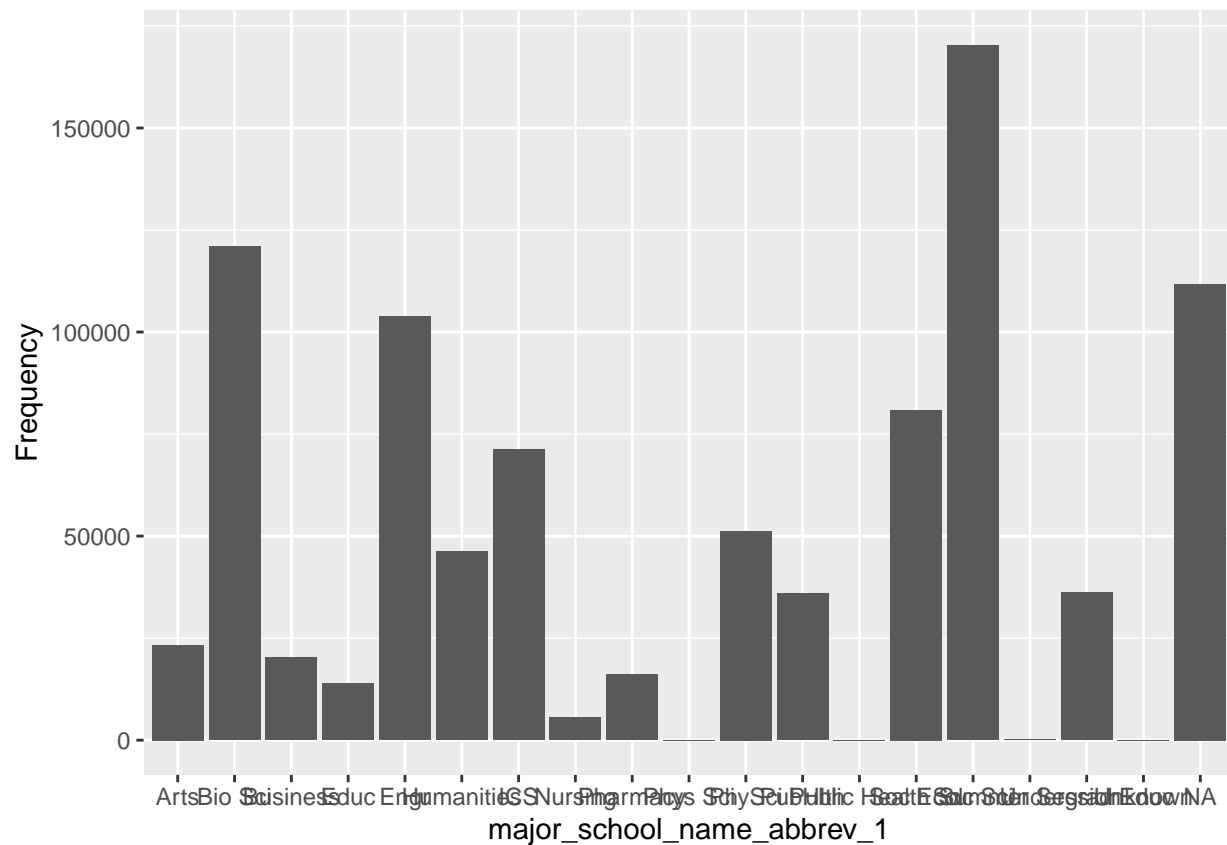
```
[1] -----
[1] Variable: major_school_code_1, type: character, requested: yes
[1] Values (16 unique): 90, 95, 08, 65, 77, ...
[1] Missing: 12.3%
  Group.1      x
1    2006 0.743810803
2    2007 0.606914730
3    2008 0.397552131
4    2009 0.119125456
5    2010 0.021535805
6    2011 0.021233934
7    2012 0.020615357
8    2013 0.020002703
9    2014 0.019496279
10   2015 0.017037388
11   2016 0.012397534
12   2017 0.007640676
13   2018 0.001349410
```



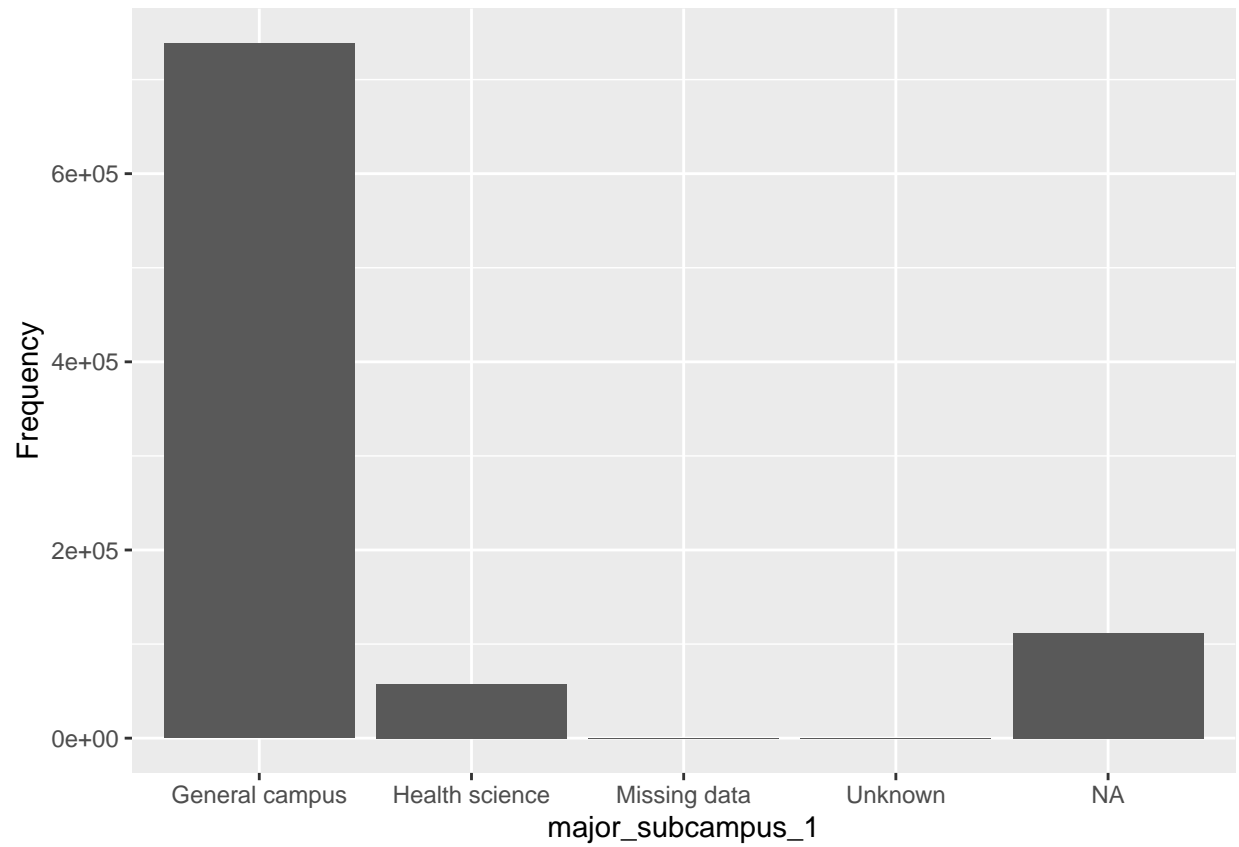
```
[1] -----
[1] Variable: major_school_name_1, type: character, requested: yes
[1] Values (17 unique): School of Nursing, School of Info & Computer Sci, Div of Undergraduate Education
[1] Missing: 12.3%
Group.1      x
1      2006 0.743810803
2      2007 0.606914730
3      2008 0.397552131
4      2009 0.119125456
5      2010 0.021535805
6      2011 0.021233934
7      2012 0.020615357
8      2013 0.020002703
9      2014 0.019496279
10     2015 0.017037388
11     2016 0.012397534
12     2017 0.007607115
13     2018 0.001349410
```



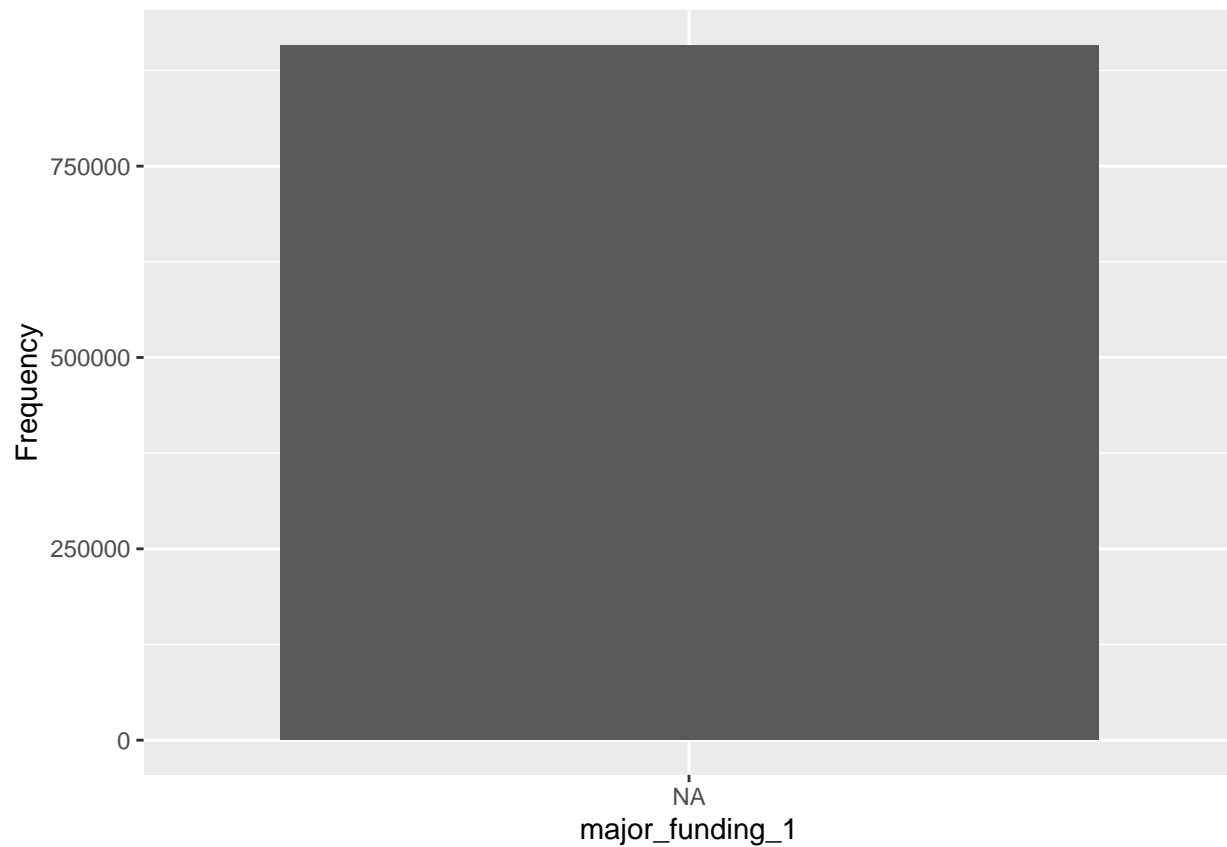
```
[1] -----
[1] Variable: major_school_name_abbrev_1, type: character, requested: yes
[1] Values (19 unique): Nursing, ICS, Undergrad Educ, Soc Sci, Engr, ...
[1] Missing: 12.3%
Group.1      x
1      2006 0.743810803
2      2007 0.606914730
3      2008 0.397552131
4      2009 0.119125456
5      2010 0.021535805
6      2011 0.021233934
7      2012 0.020615357
8      2013 0.020002703
9      2014 0.019496279
10     2015 0.017037388
11     2016 0.012397534
12     2017 0.007607115
13     2018 0.001349410
```



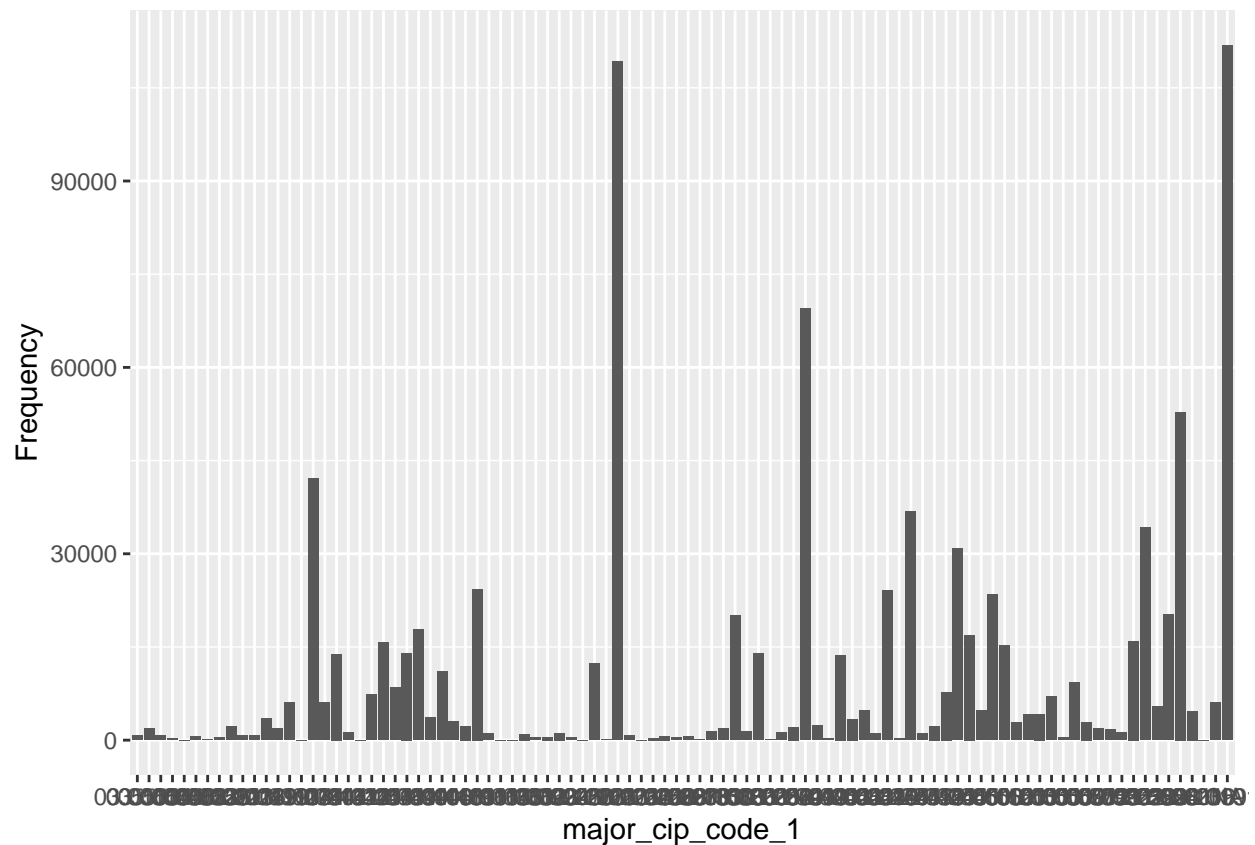
```
[1] -----
[1] Variable: major_subcampus_1, type: character, requested: yes
[1] Values (5 unique): Health science, General campus, NA, Unknown, Missing data
[1] Missing: 12.3%
  Group.1      x
1    2006 0.743810803
2    2007 0.606914730
3    2008 0.397552131
4    2009 0.119125456
5    2010 0.021535805
6    2011 0.021233934
7    2012 0.020615357
8    2013 0.020002703
9    2014 0.019496279
10   2015 0.017037388
11   2016 0.012397534
12   2017 0.007607115
13   2018 0.001349410
```



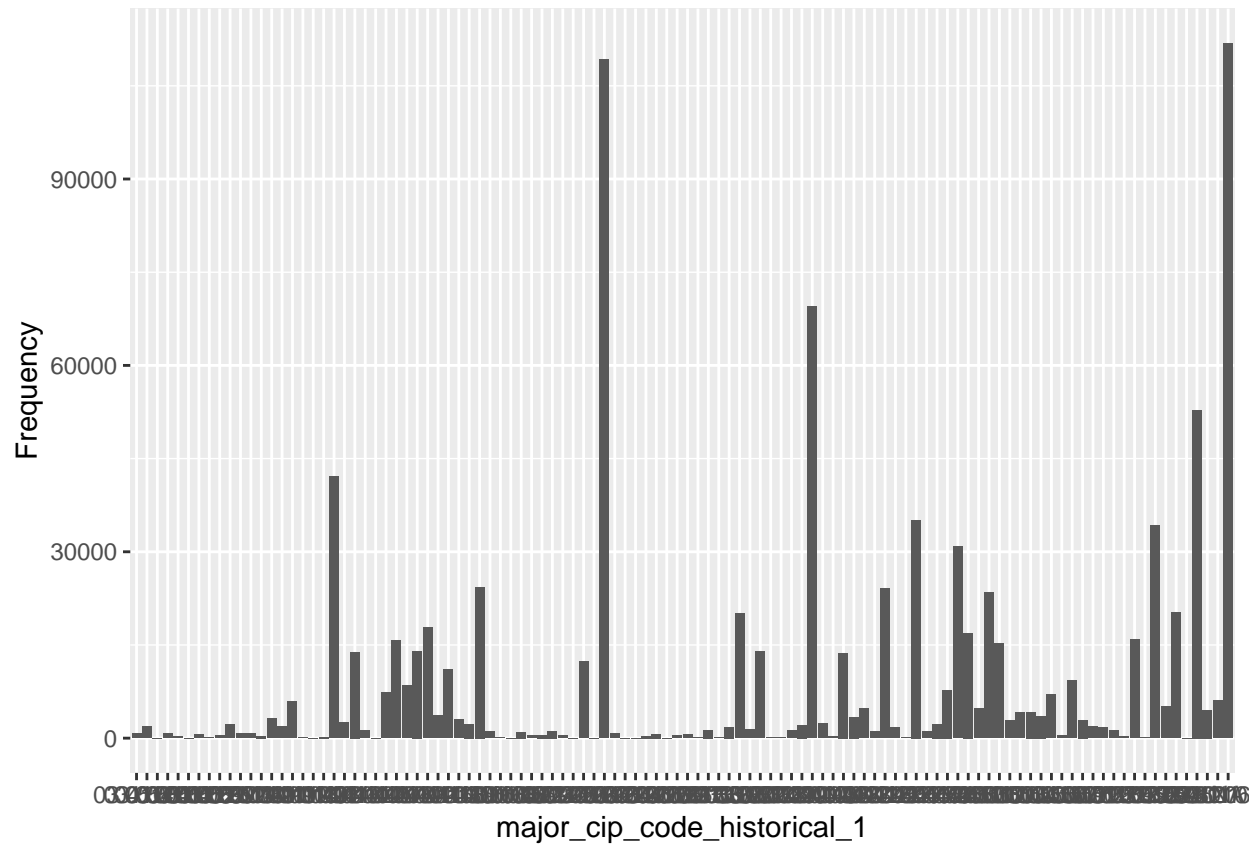
```
[1] -----  
[1] Variable: major_funding_1, type: logical, requested: yes  
[1] Values (1 unique): NA  
[1] Missing: 100%
```



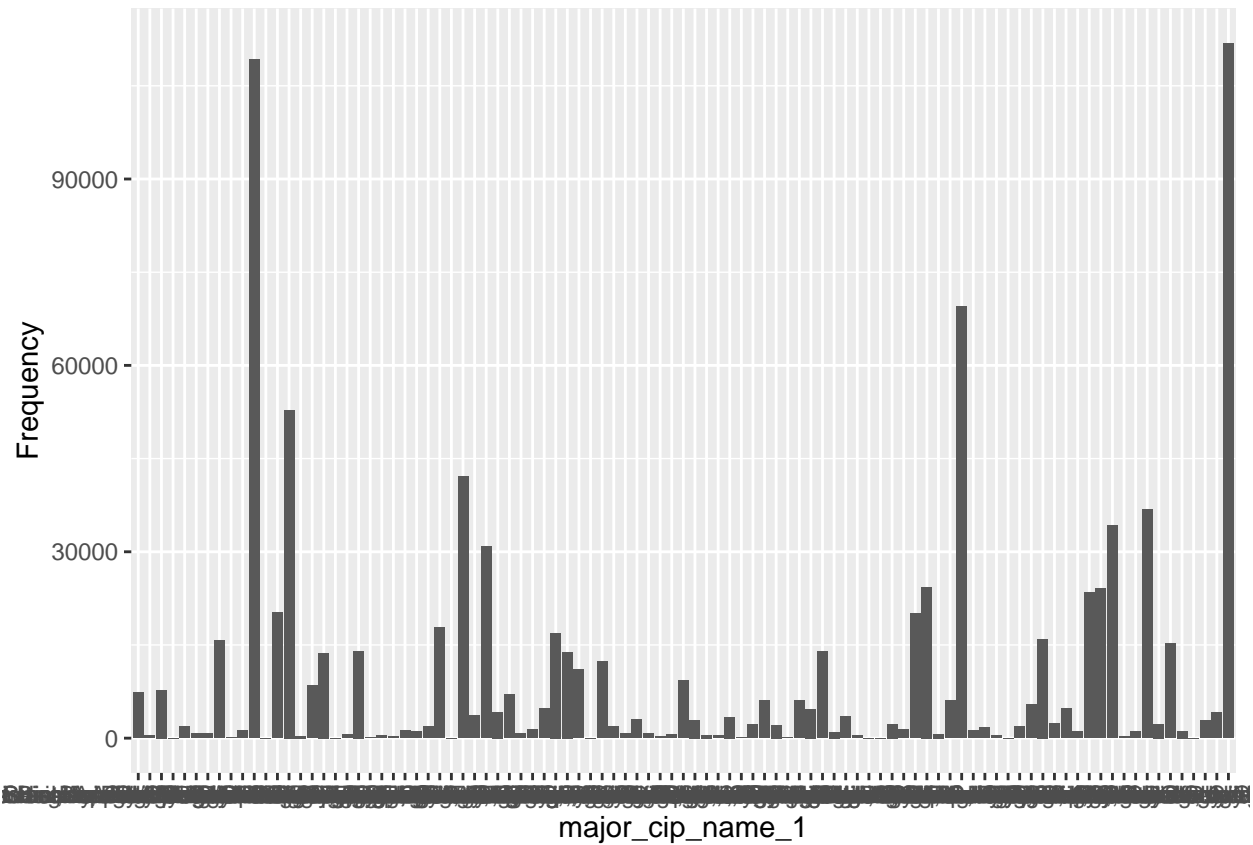
```
[1] -----
[1] Variable: major_cip_code_1, type: character, requested: yes
[1] Values (94 unique): 51.3808, 11.0701, 30.9999, 45.1101, 05.0203, ...
[1] Missing: 12.3%
  Group.1      x
1    2006 0.743810803
2    2007 0.606914730
3    2008 0.397567762
4    2009 0.119125456
5    2010 0.021535805
6    2011 0.021248359
7    2012 0.020630135
8    2013 0.020029734
9    2014 0.019523245
10   2015 0.017076826
11   2016 0.012623354
12   2017 0.007853227
13   2018 0.001799213
```

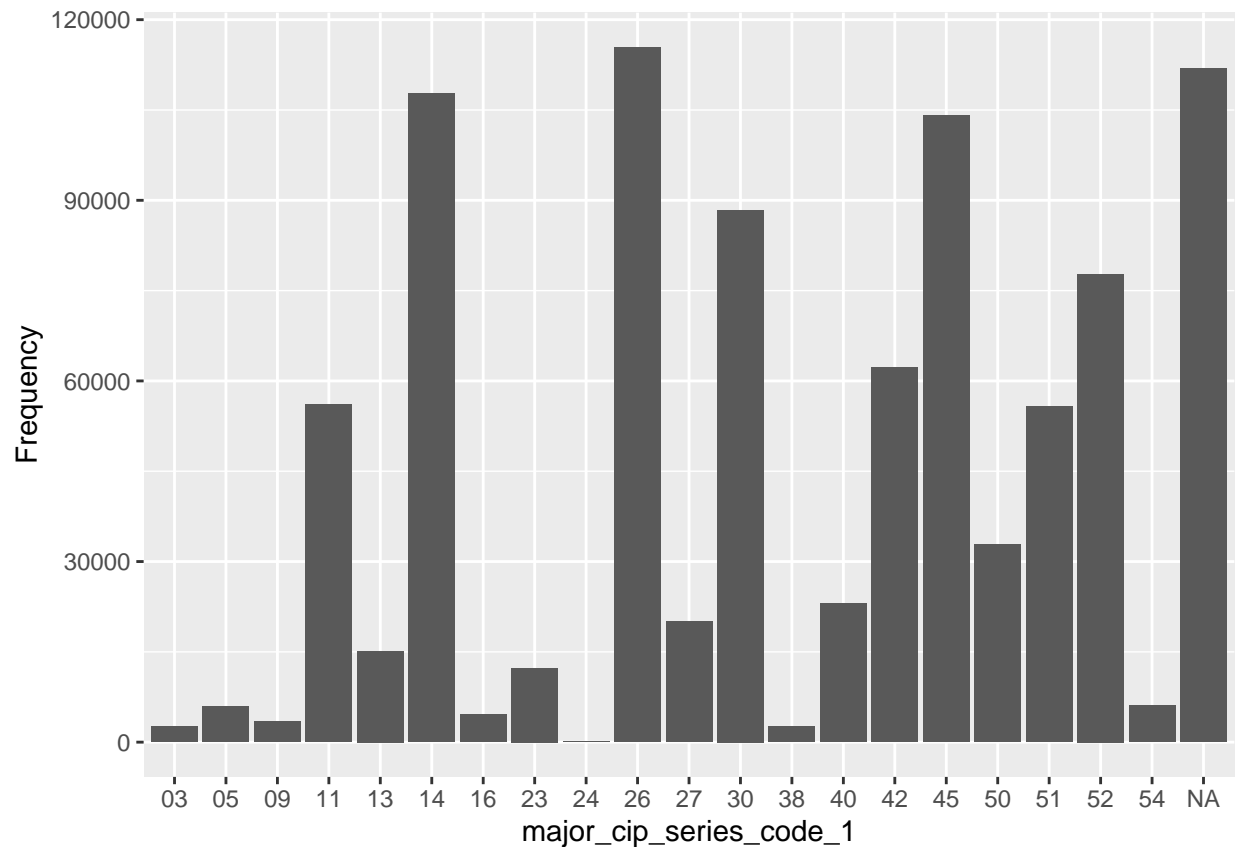
```
[1] -----
[1] Variable: major_cip_code_historical_1, type: character, requested: yes
[1] Values (106 unique): 51.3808, 11.0701, 30.9999, 45.1101, 05.0203, ...
[1] Missing: 12.3%
Group.1      x
1      2006 0.743810803
2      2007 0.606914730
3      2008 0.397567762
4      2009 0.119125456
5      2010 0.021535805
6      2011 0.021248359
7      2012 0.020630135
8      2013 0.020029734
9      2014 0.019523245
10     2015 0.017076826
11     2016 0.012623354
12     2017 0.007853227
13     2018 0.001799213
```



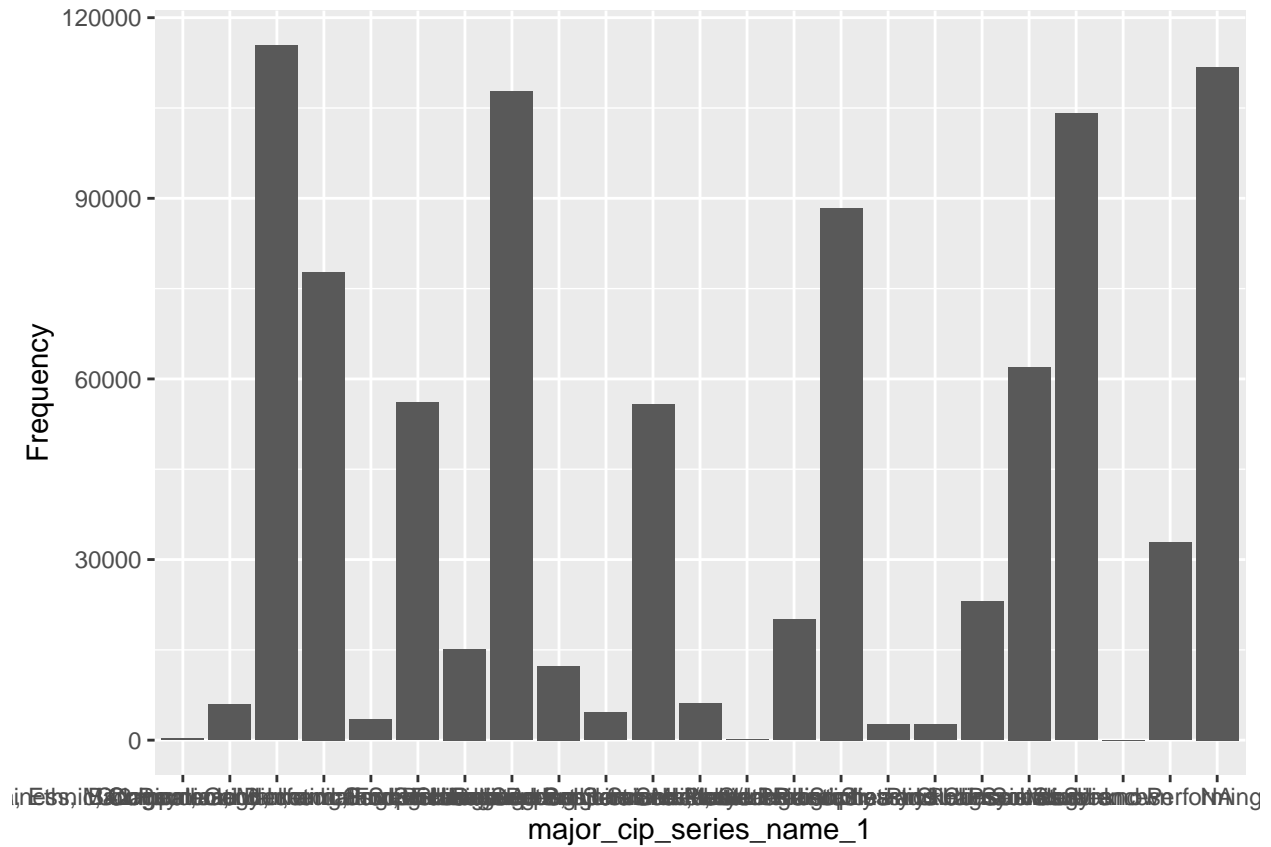
```
[1] -----
[1] Variable: major_cip_name_1, type: character, requested: yes
[1] Values (95 unique): Nursing Science, Computer Science, Multi-/Interdisciplinary Studies, Other, Soc
[1] Missing: 12.3%
  Group.1      x
1    2006 0.743810803
2    2007 0.606914730
3    2008 0.397567762
4    2009 0.119125456
5    2010 0.021535805
6    2011 0.021248359
7    2012 0.020630135
8    2013 0.020029734
9    2014 0.019523245
10   2015 0.017076826
11   2016 0.012623354
12   2017 0.007819667
13   2018 0.001799213
```



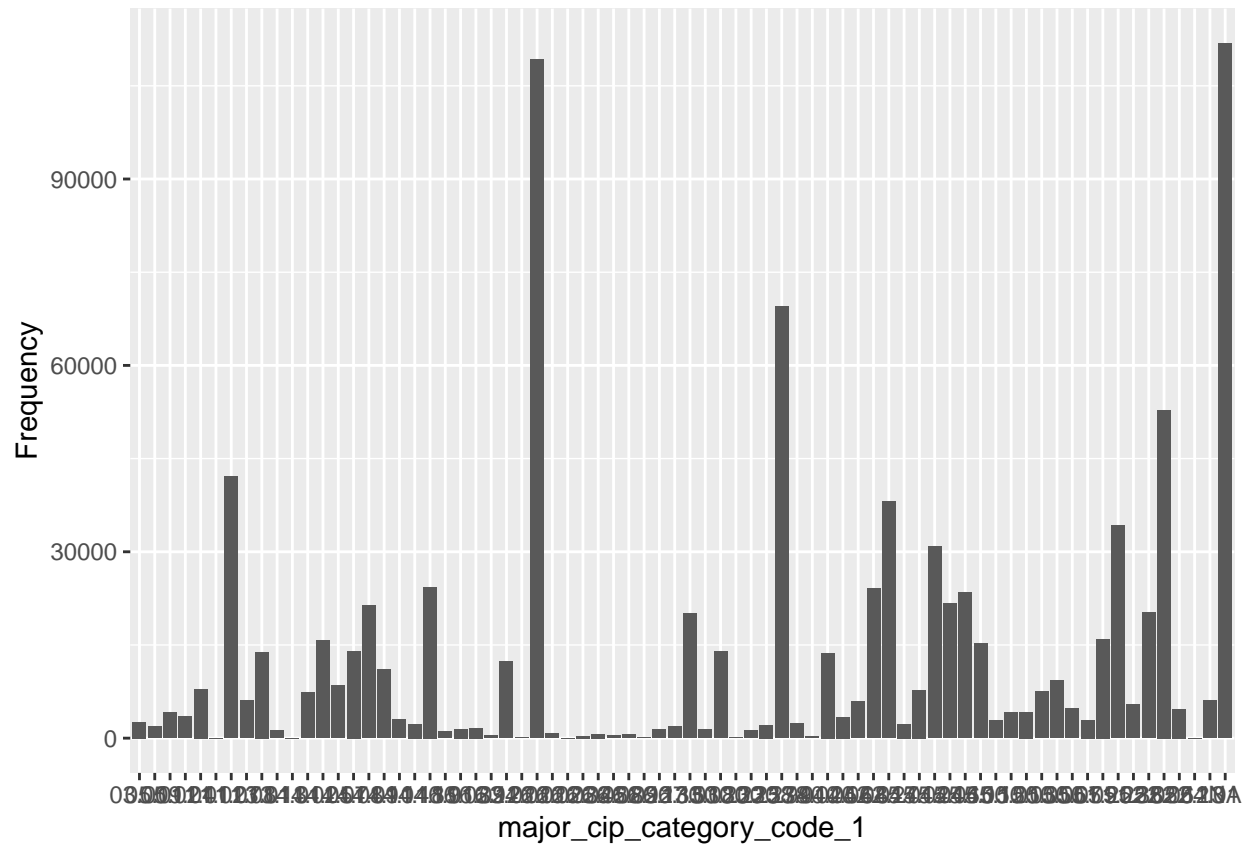
```
[1] -----
[1] Variable: major_cip_series_code_1, type: character, requested: yes
[1] Values (21 unique): 51, 11, 30, 45, 05, ...
[1] Missing: 12.3%
  Group.1      x
1    2006 0.743810803
2    2007 0.606914730
3    2008 0.397567762
4    2009 0.119125456
5    2010 0.021535805
6    2011 0.021248359
7    2012 0.020630135
8    2013 0.020029734
9    2014 0.019523245
10   2015 0.017076826
11   2016 0.012623354
12   2017 0.007853227
13   2018 0.001799213
```



```
[1] -----
[1] Variable: major_cip_series_name_1, type: character, requested: yes
[1] Values (23 unique): Health Professions and Related Programs, Computer and Information Sciences and S
[1] Missing: 12.3%
  Group.1      x
1    2006 0.743810803
2    2007 0.606914730
3    2008 0.397567762
4    2009 0.119125456
5    2010 0.021535805
6    2011 0.021248359
7    2012 0.020630135
8    2013 0.020029734
9    2014 0.019523245
10   2015 0.017076826
11   2016 0.012623354
12   2017 0.007819667
13   2018 0.001799213
```

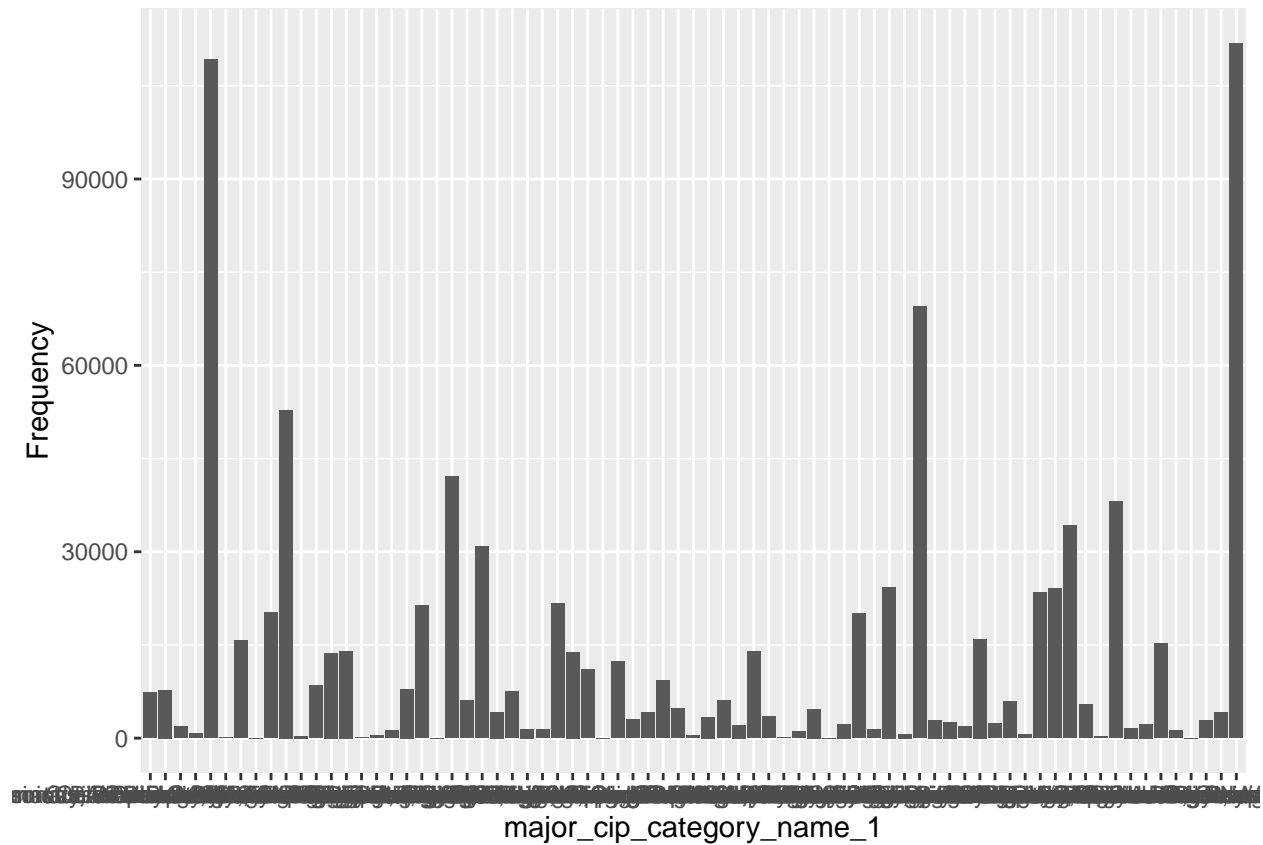


```
[1] -----
[1] Variable: major_cip_category_code_1, type: character, requested: yes
[1] Values (72 unique): 51.38, 11.07, 30.99, 45.11, 05.02, ...
[1] Missing: 12.3%
Group.1      x
1      2006 0.743810803
2      2007 0.606914730
3      2008 0.397567762
4      2009 0.119125456
5      2010 0.021535805
6      2011 0.021248359
7      2012 0.020630135
8      2013 0.020029734
9      2014 0.019523245
10     2015 0.017076826
11     2016 0.012623354
12     2017 0.007853227
13     2018 0.001799213
```



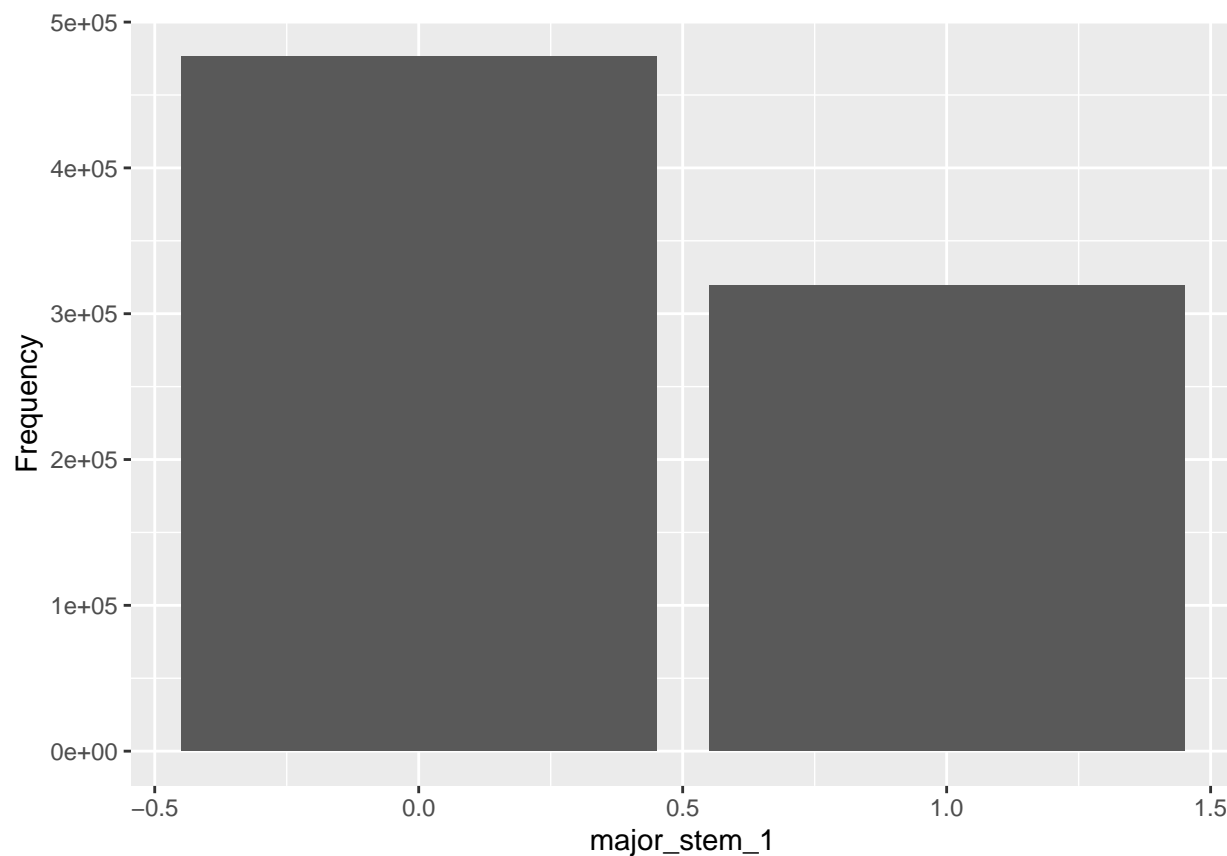
```
[1] -----
[1] Variable: major_cip_category_name_1, type: character, requested: yes
[1] Values (73 unique): Registered Nursing, Nursing Administration, Nursing Research and Clinical Nursin
[1] Missing: 12.3%

  Group.1      x
1    2006 0.743810803
2    2007 0.606914730
3    2008 0.397567762
4    2009 0.119125456
5    2010 0.021535805
6    2011 0.021248359
7    2012 0.020630135
8    2013 0.020029734
9    2014 0.019523245
10   2015 0.017076826
11   2016 0.012623354
12   2017 0.007819667
13   2018 0.001799213
```



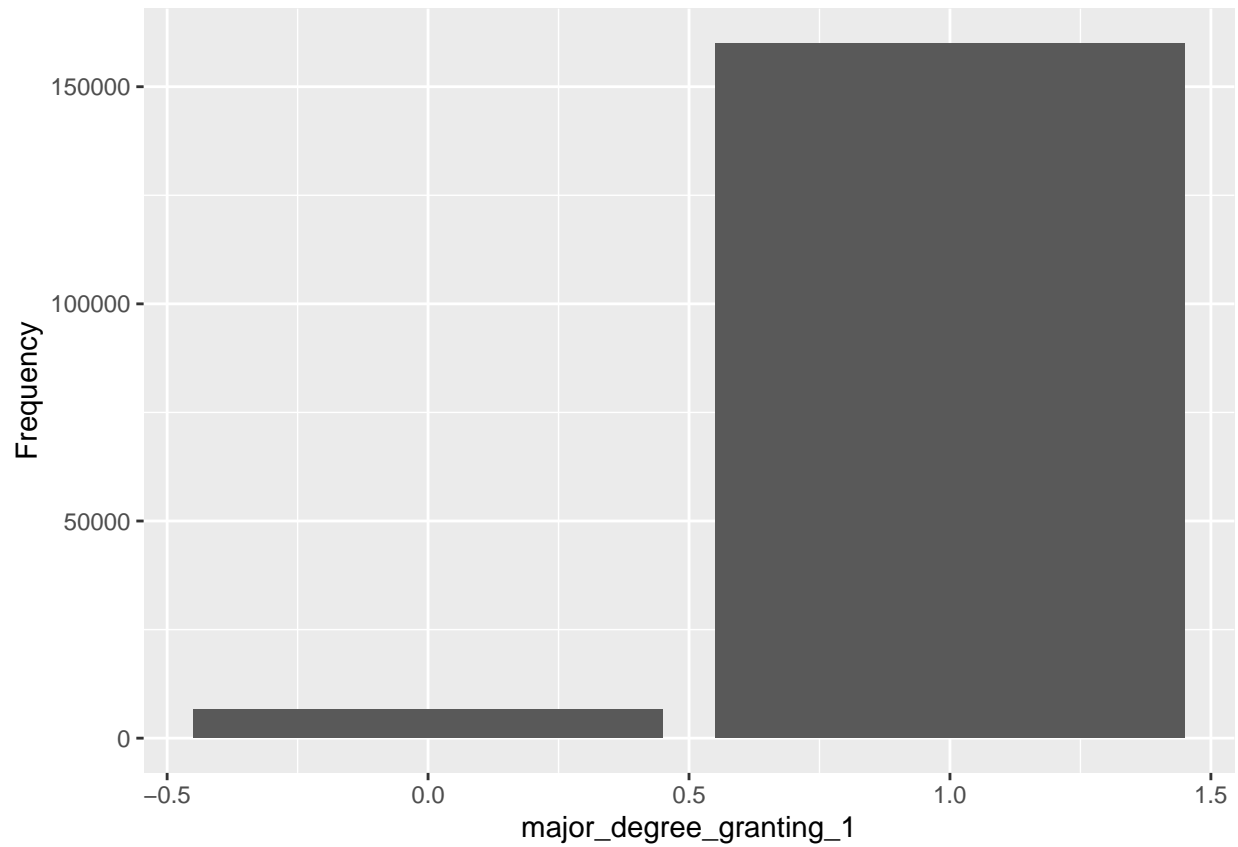
```
[1] -----
[1] Variable: major_stem_1, type: numeric, requested: yes
[1] Values (3 unique): 0, 1, NA
[1] Missing: 12.3%
  Group.1      x
1   2006 0.743810803
2   2007 0.606914730
3   2008 0.397552131
4   2009 0.119125456
5   2010 0.021535805
6   2011 0.021233934
7   2012 0.020615357
8   2013 0.020002703
9   2014 0.019496279
10  2015 0.017037388
11  2016 0.012397534
12  2017 0.007640676
13  2018 0.001349410
```

Warning: Removed 111753 rows containing non-finite values ('stat_count()').

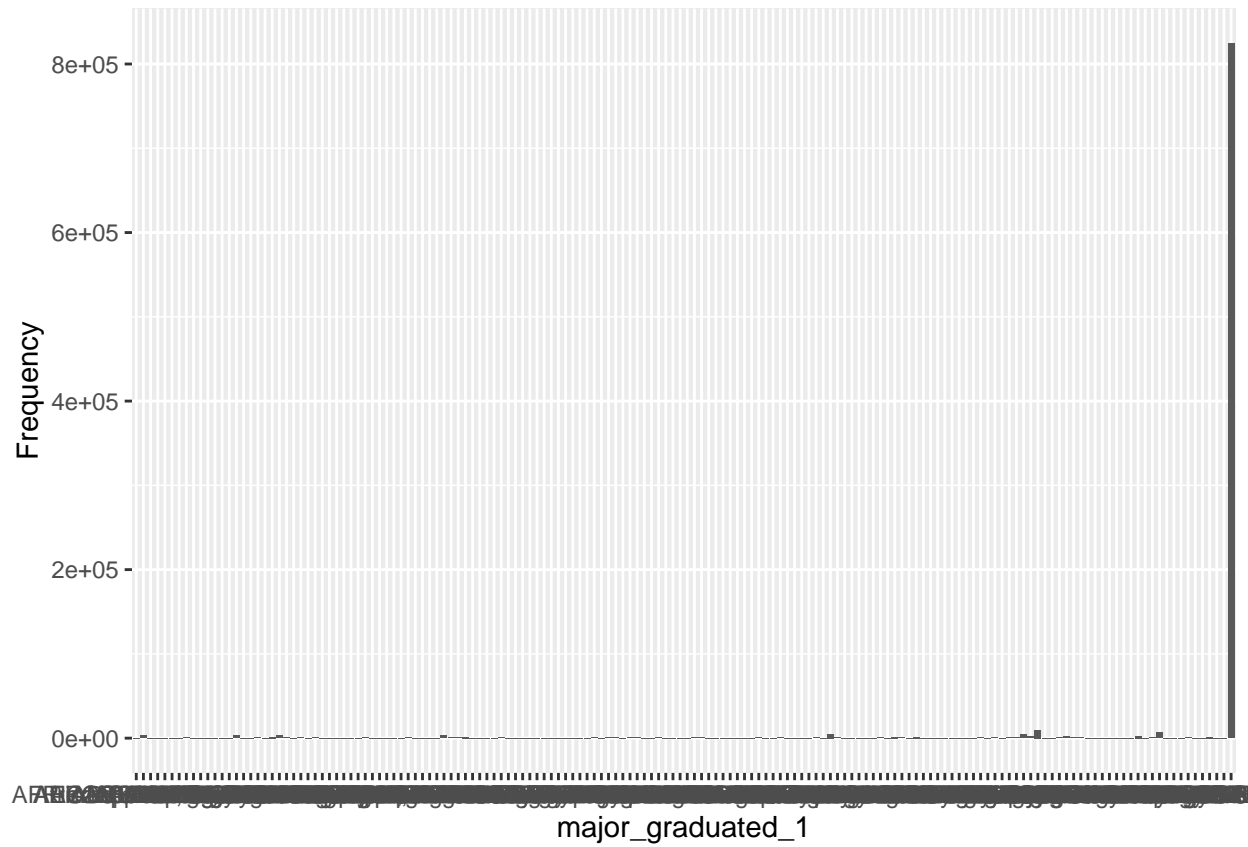


```
[1] -----
[1] Variable: major_degree_granting_1, type: numeric, requested: yes
[1] Values (3 unique): NA, 1, 0
[1] Missing: 81.6%
  Group.1      x
1    2006 0.9989164
2    2007 0.9995310
3    2008 0.9995311
4    2009 0.9992562
5    2010 0.9993125
6    2011 0.9980093
7    2012 0.9972217
8    2013 0.9944993
9    2014 0.9769173
10   2015 0.8583504
11   2016 0.6640472
12   2017 0.4191185
13   2018 0.1058037
```

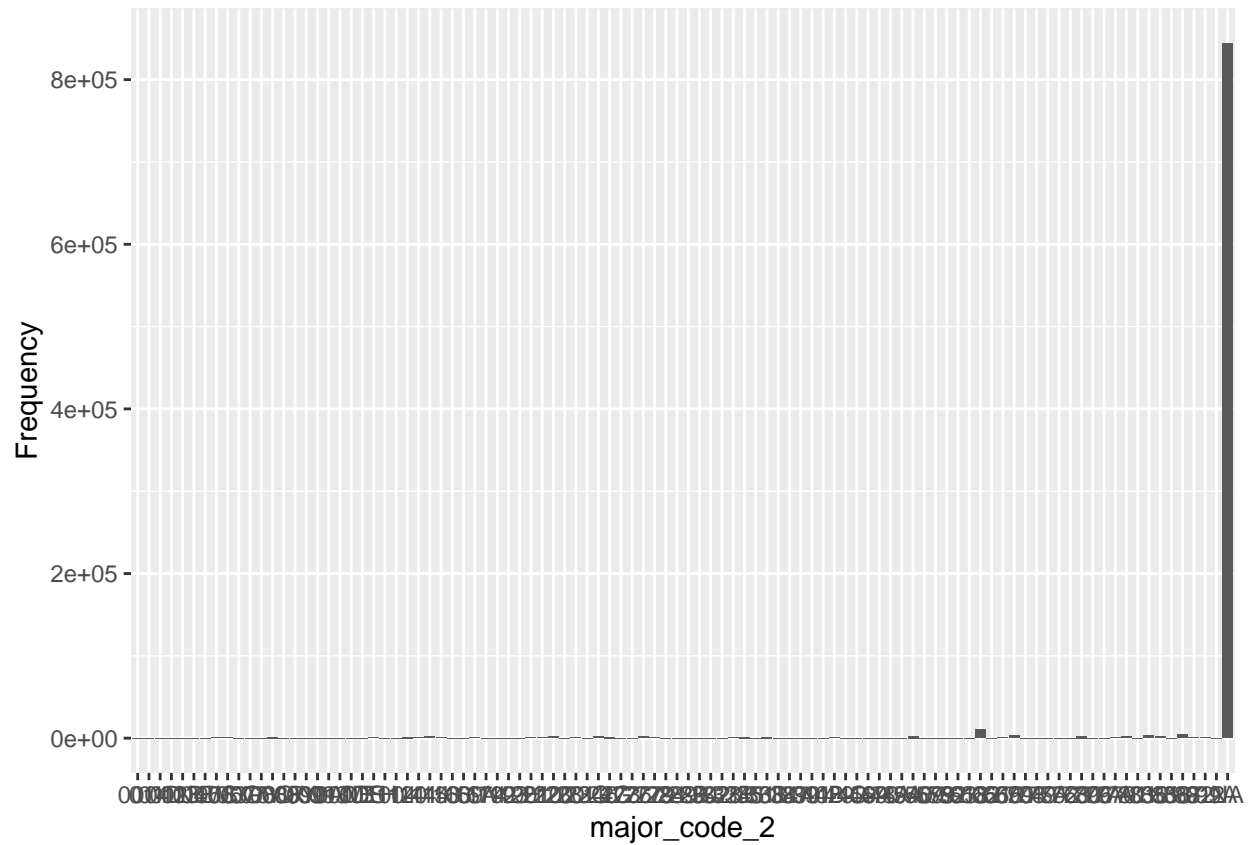
Warning: Removed 741177 rows containing non-finite values ('stat_count()').



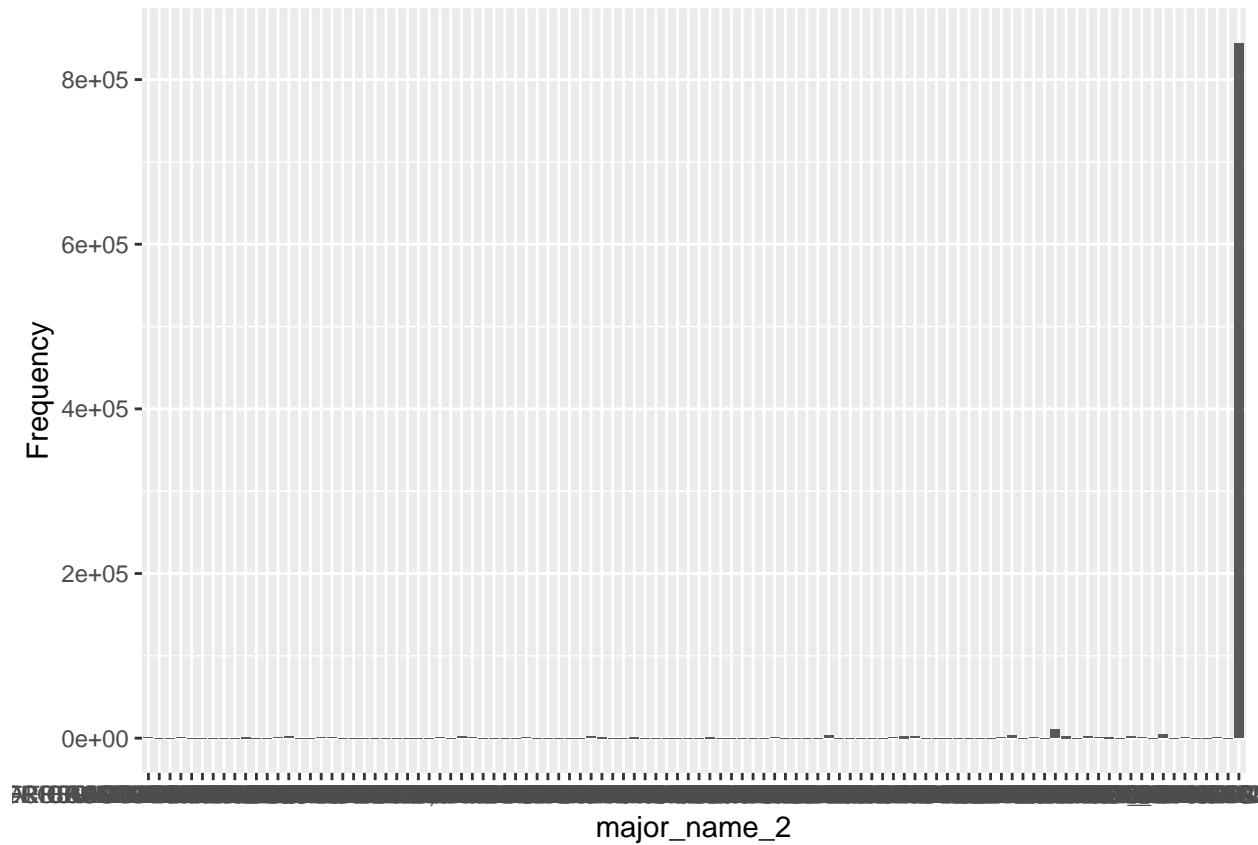
```
[1] -----  
[1] Variable: major_graduated_1, type: character, requested: yes  
[1] Values (154 unique): NA, Computer Science, SOCIOLOGY, Business Economics, Civil Engineering, ...  
[1] Missing: 90.8%  
[1] Most missing: 2009 94.9%, Least missing: 2014 88.5%
```



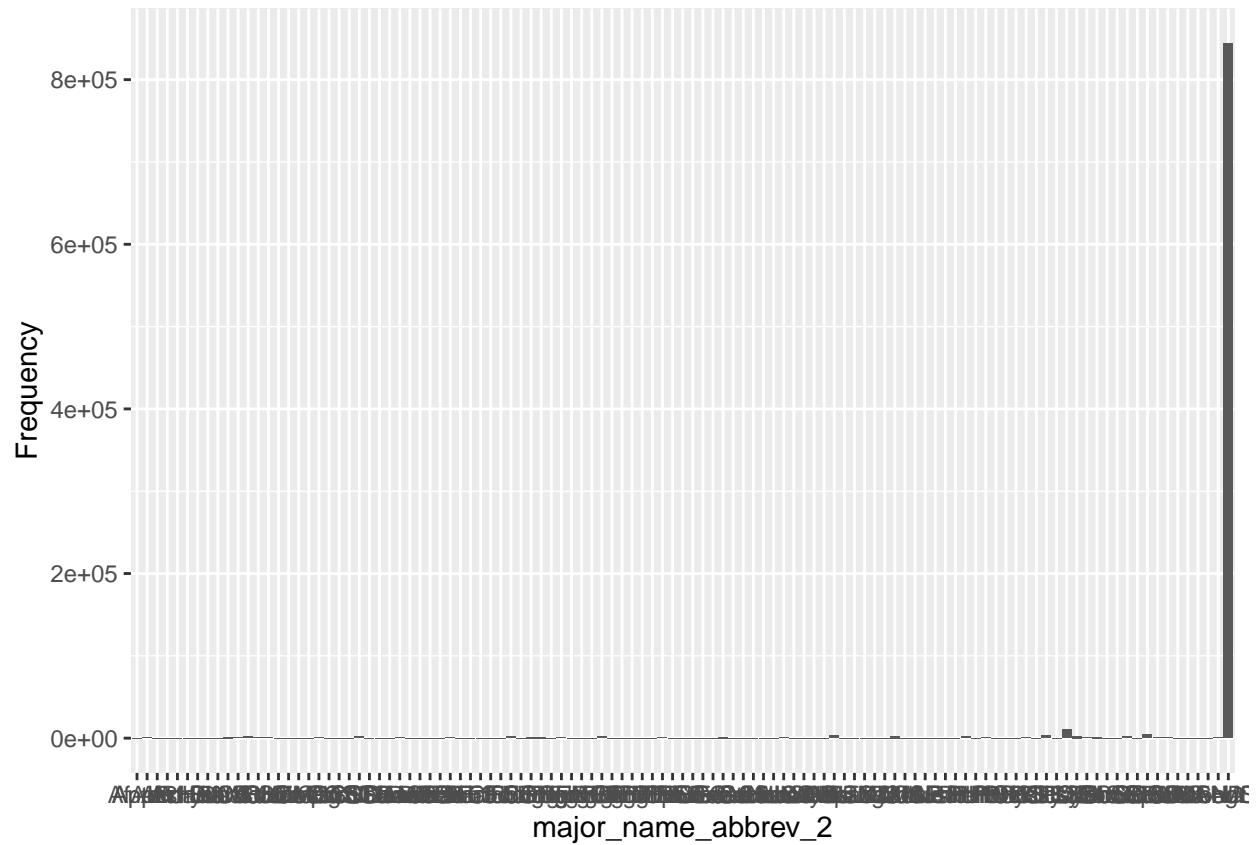
```
[1] -----
[1] Variable: major_code_2, type: character, requested: yes
[1] Values (98 unique): NA, 867, 345, 7A1, 192, ...
[1] Missing: 93%
[1] Most missing: 2006 95.9%, Least missing: 2015 92.3%
```



```
[1] -----
[1] Variable: major_name_2, type: character, requested: yes
[1] Values (102 unique): NA, SOCIOLOGY, ENGLISH, PUBLIC HEALTH POLICY, COMPARATIVE LITERATURE, ...
[1] Missing: 93%
[1] Most missing: 2006 95.9%, Least missing: 2015 92.3%
```

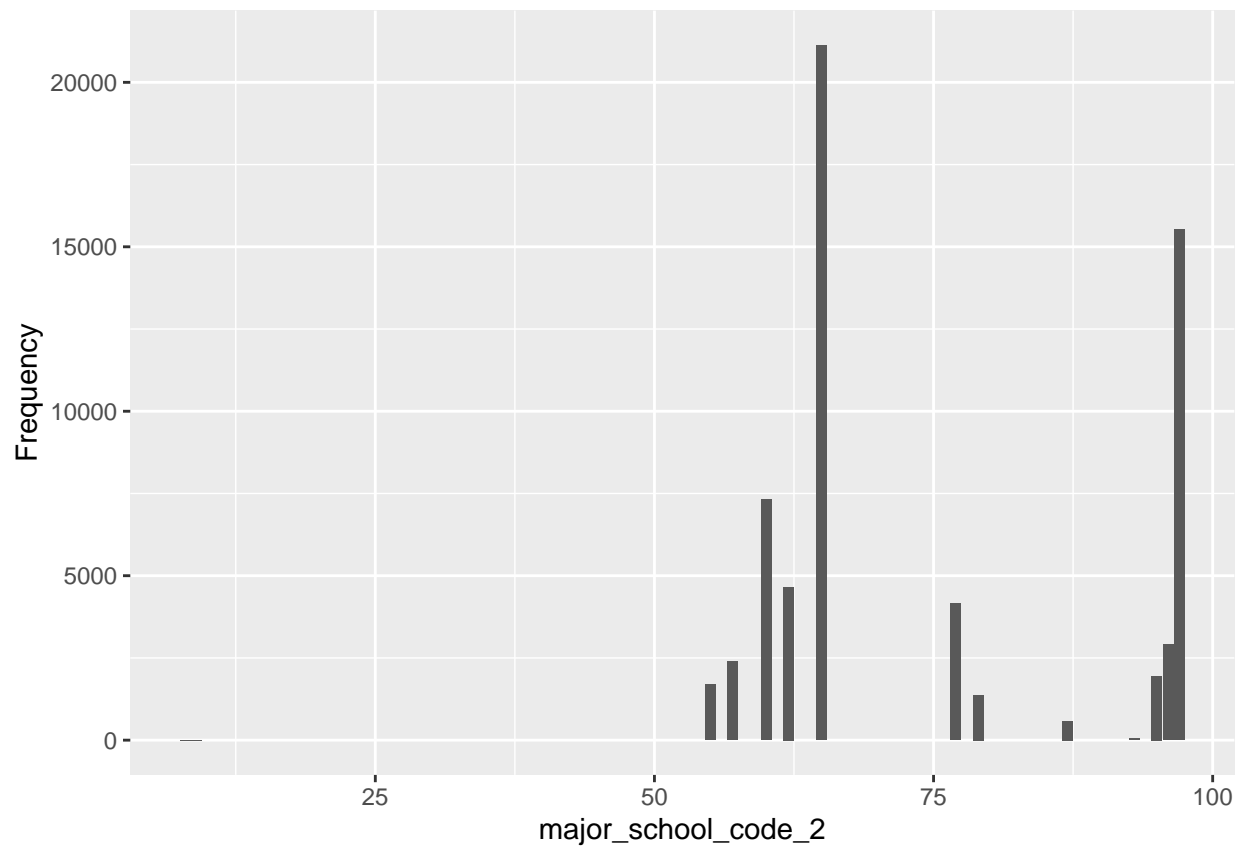


```
[1] -----
[1] Variable: major_name_abbrev_2, type: character, requested: yes
[1] Values (109 unique): NA, Sociol, English, P H Pol, Com Lit, ...
[1] Missing: 93%
[1] Most missing: 2006 95.9%, Least missing: 2015 92.3%
```

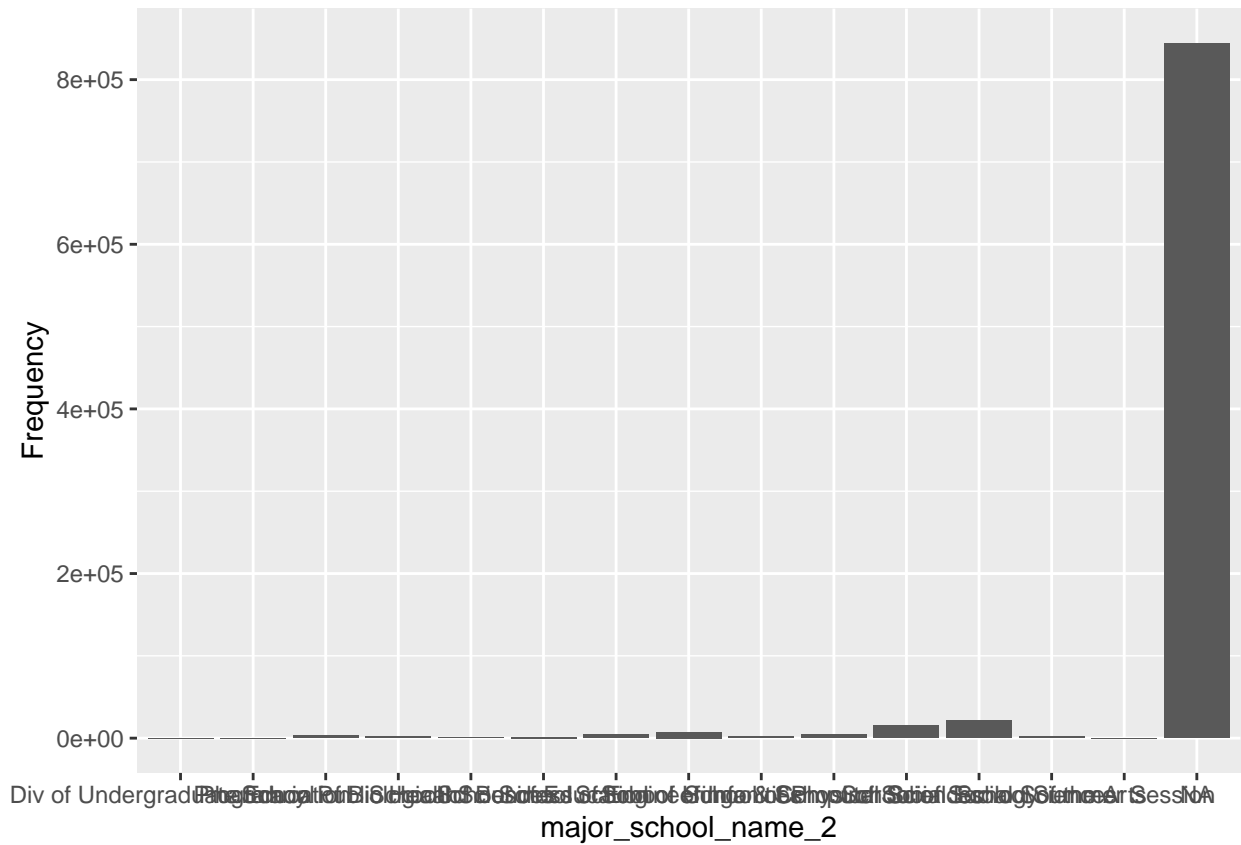


```
[1] -----
[1] Variable: major_school_code_2, type: numeric, requested: yes
[1] Values (15 unique): NA, 65, 60, 96, 87, ...
[1] Missing: 93%
[1] Most missing: 2006 95.9%, Least missing: 2015 92.3%
```

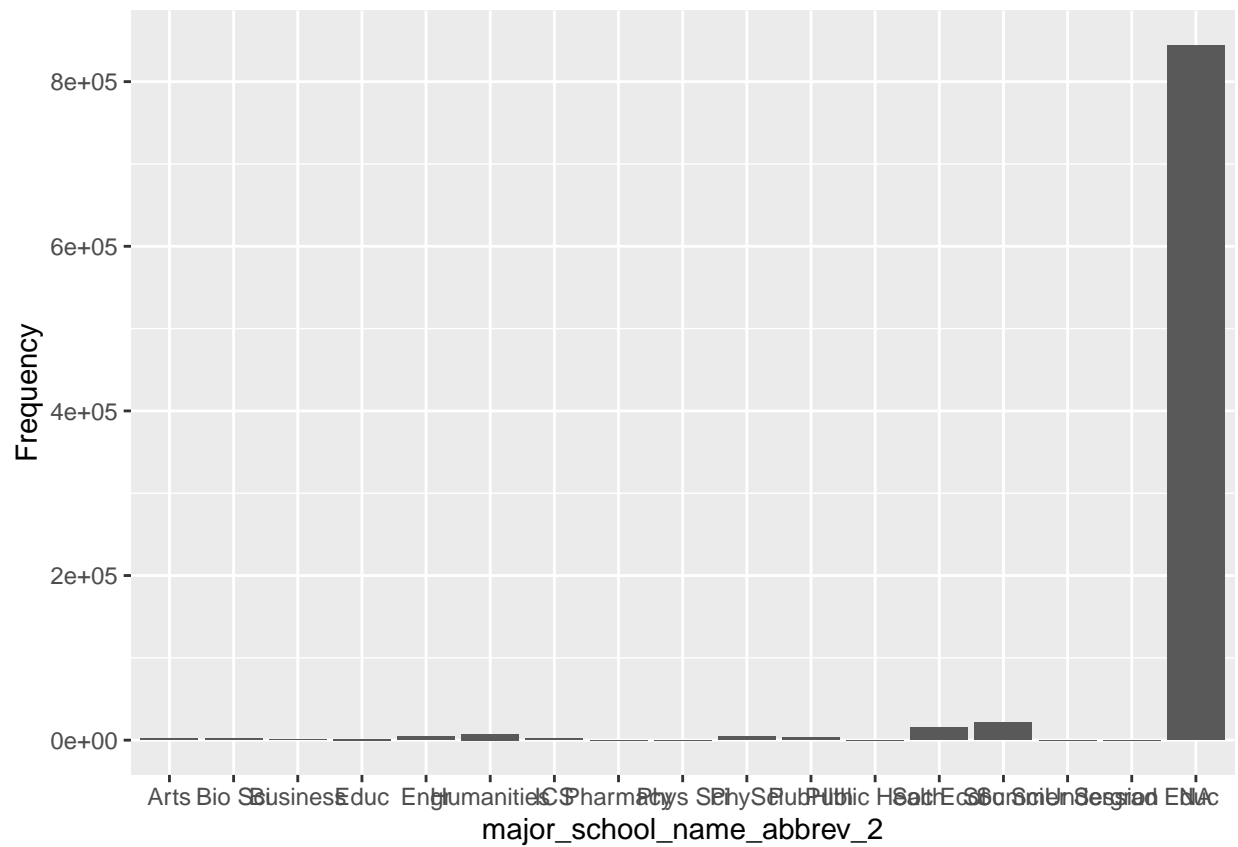
Warning: Removed 844109 rows containing non-finite values ('stat_count()').



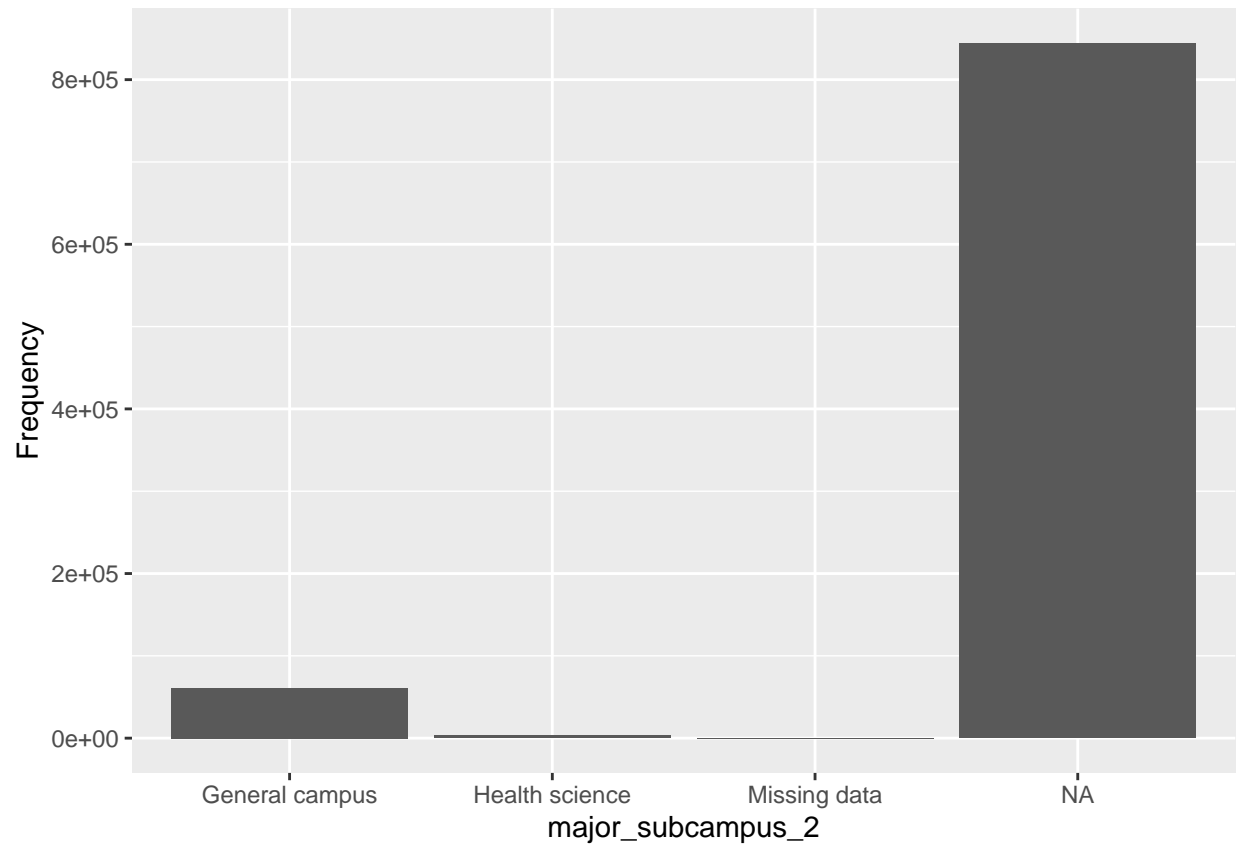
```
[1] -----
[1] Variable: major_school_name_2, type: character, requested: yes
[1] Values (15 unique): NA, School of Social Sciences, School of Humanities, Program in Public Health, S
[1] Missing: 93%
[1] Most missing: 2006 95.9%, Least missing: 2015 92.3%
```



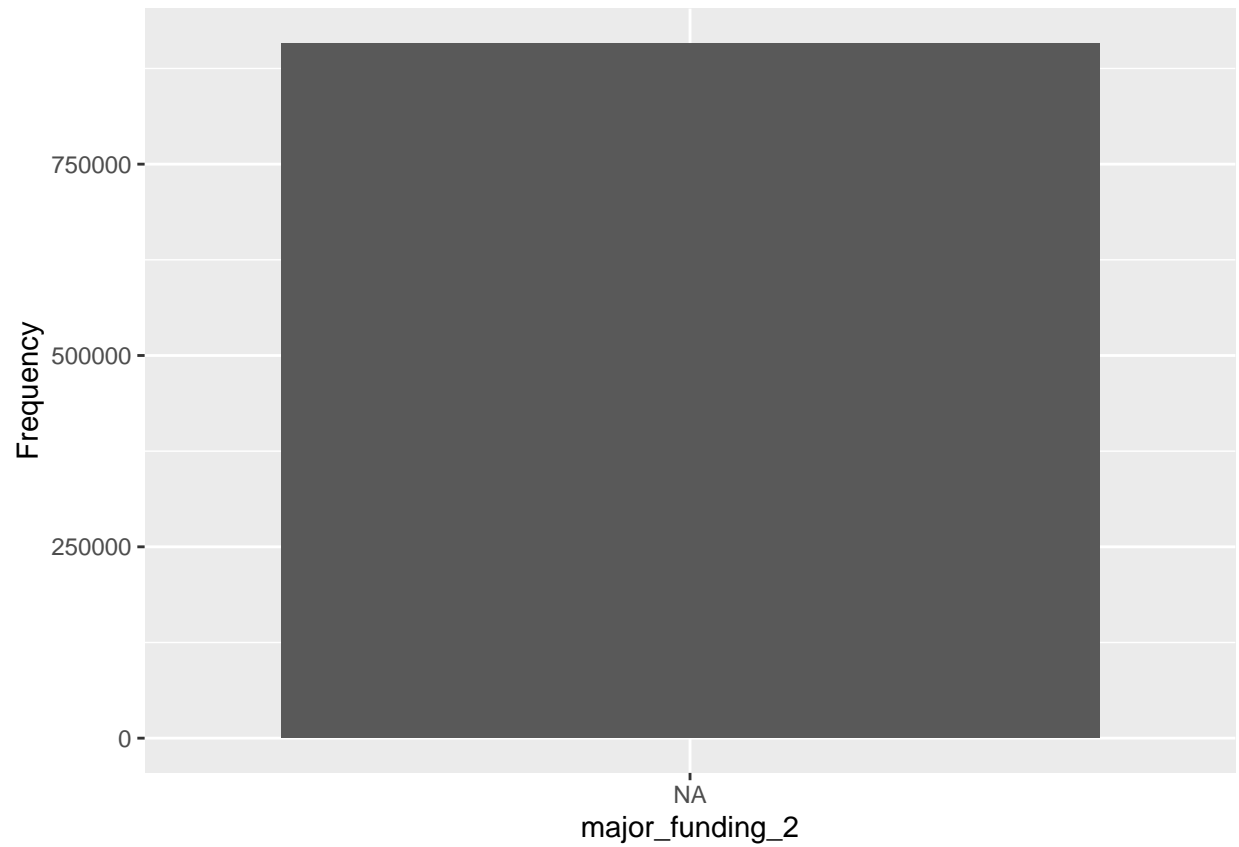
```
[1] -----
[1] Variable: major_school_name_abbrev_2, type: character, requested: yes
[1] Values (17 unique): NA, Soc Sci, Humanities, PubHlth, Business, ...
[1] Missing: 93%
[1] Most missing: 2006 95.9%, Least missing: 2015 92.3%
```



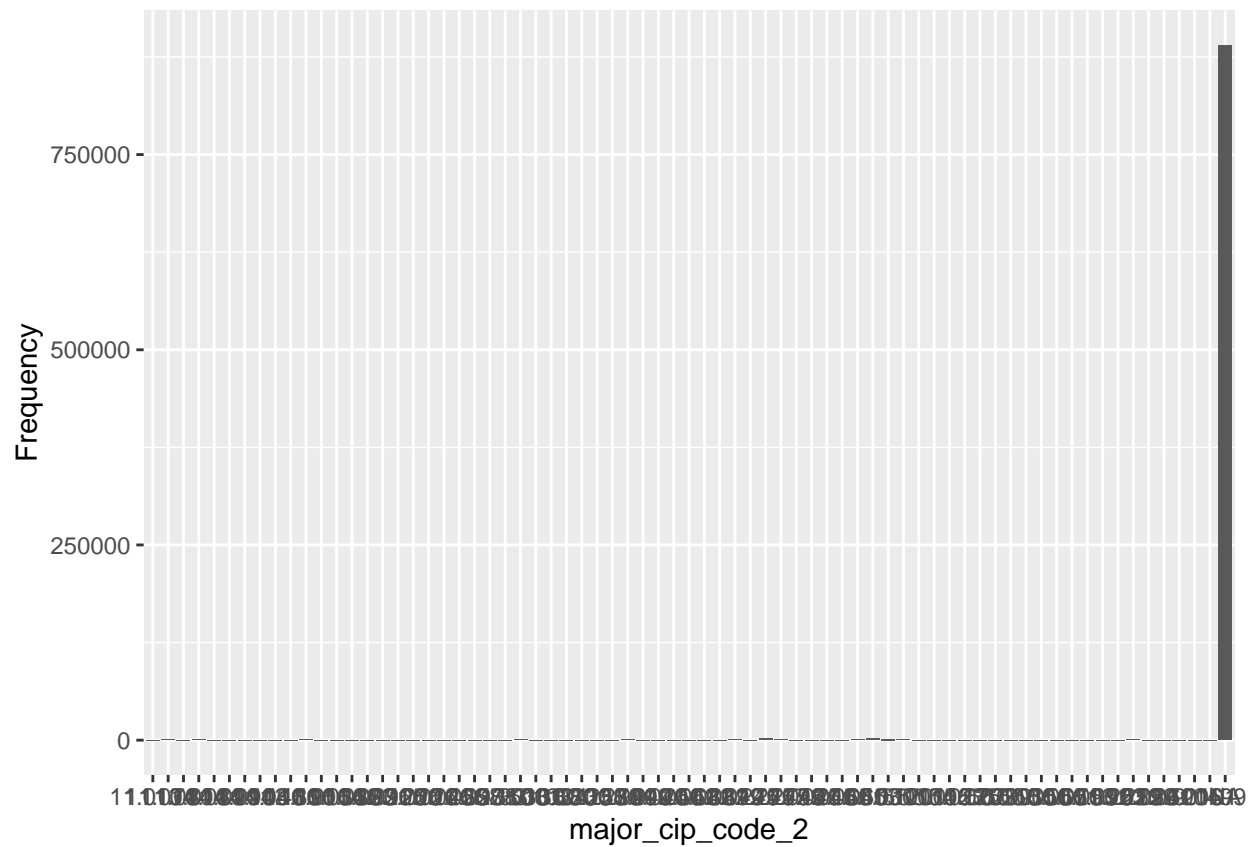
```
[1] -----
[1] Variable: major_subcampus_2, type: character, requested: yes
[1] Values (4 unique): NA, General campus, Health science, Missing data
[1] Missing: 93%
[1] Most missing: 2006 95.9%, Least missing: 2015 92.3%
```

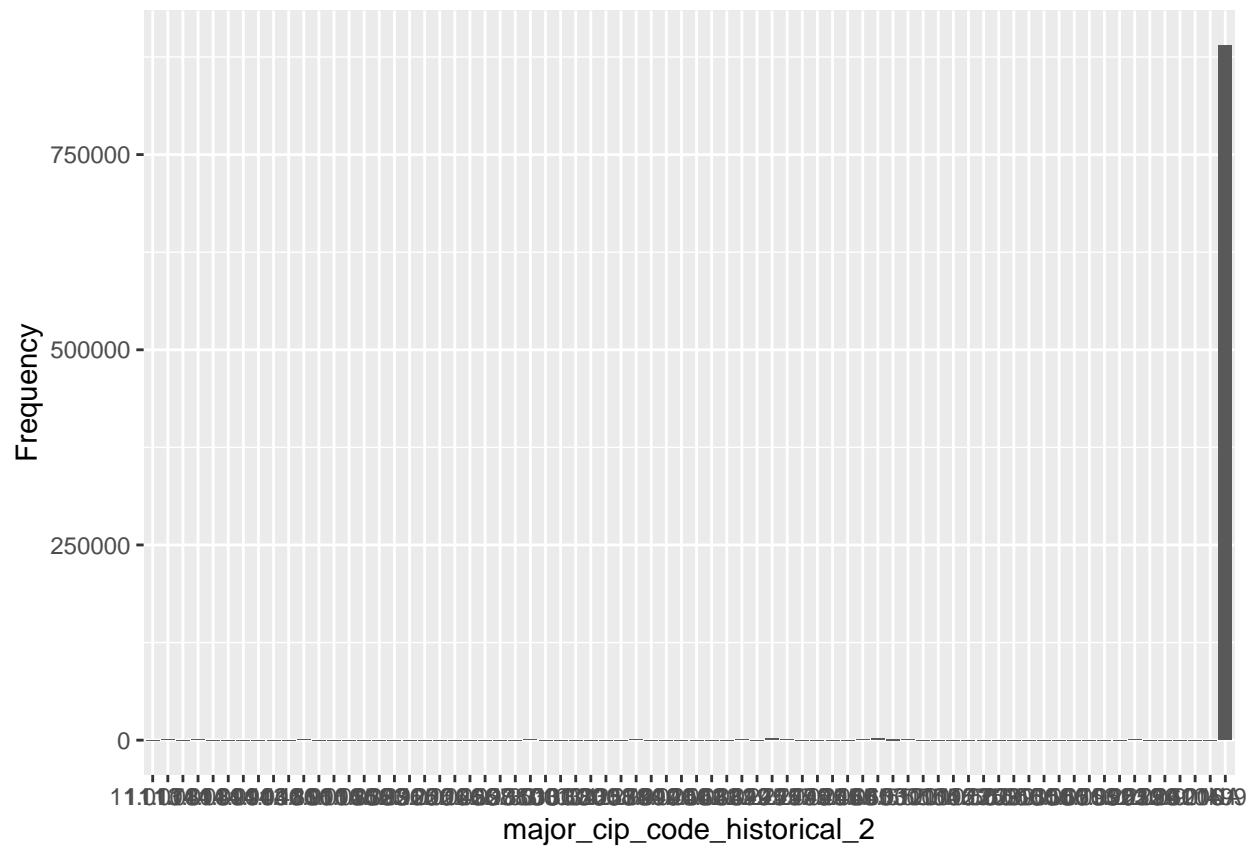
```
[1] -----  
[1] Variable: major_funding_2, type: logical, requested: yes  
[1] Values (1 unique): NA  
[1] Missing: 100%
```



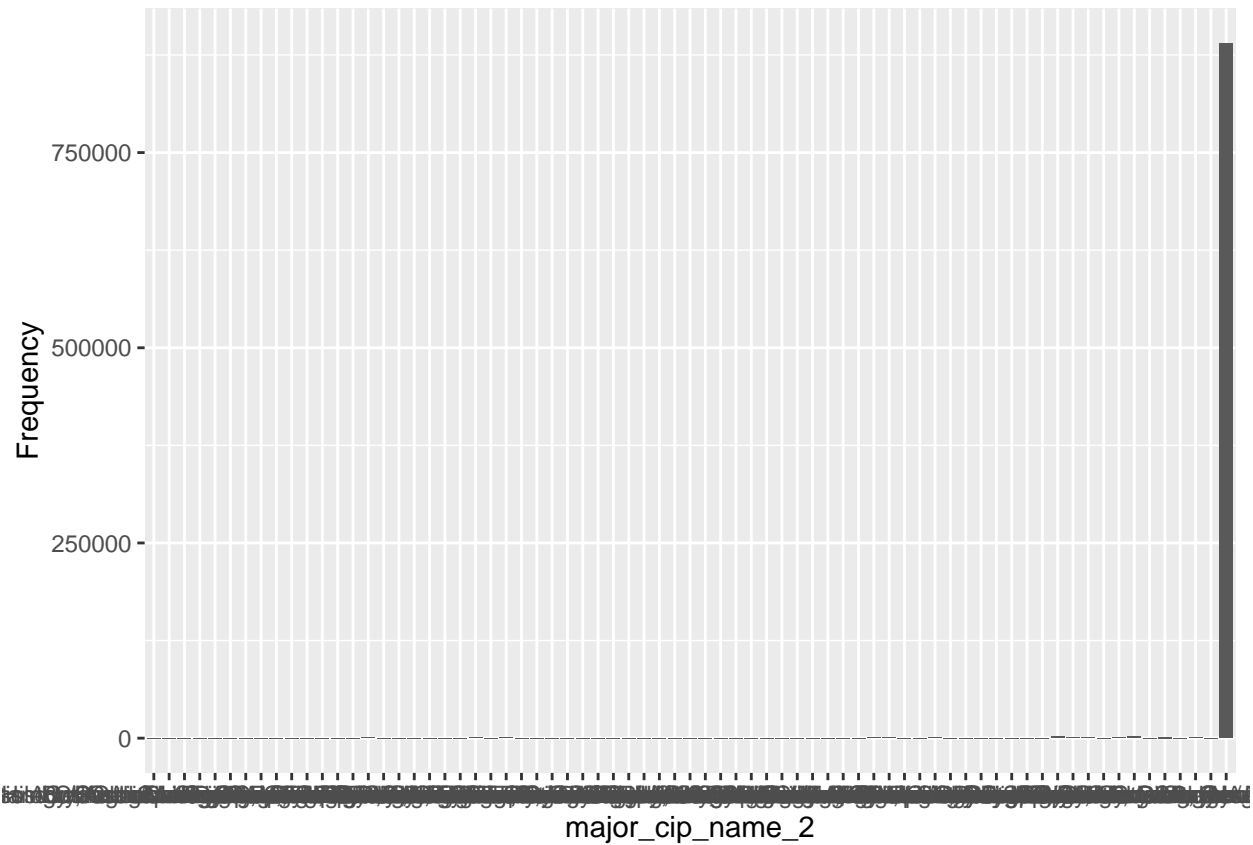
```
[1] -----  
[1] Variable: major_cip_code_2, type: character, requested: yes  
[1] Values (71 unique): NA, 45.1101, 51.2299, 23.0101, 11.0701, ...  
[1] Missing: 98%  
[1] Most missing: 2006 100%, Least missing: 2018 93.1%
```



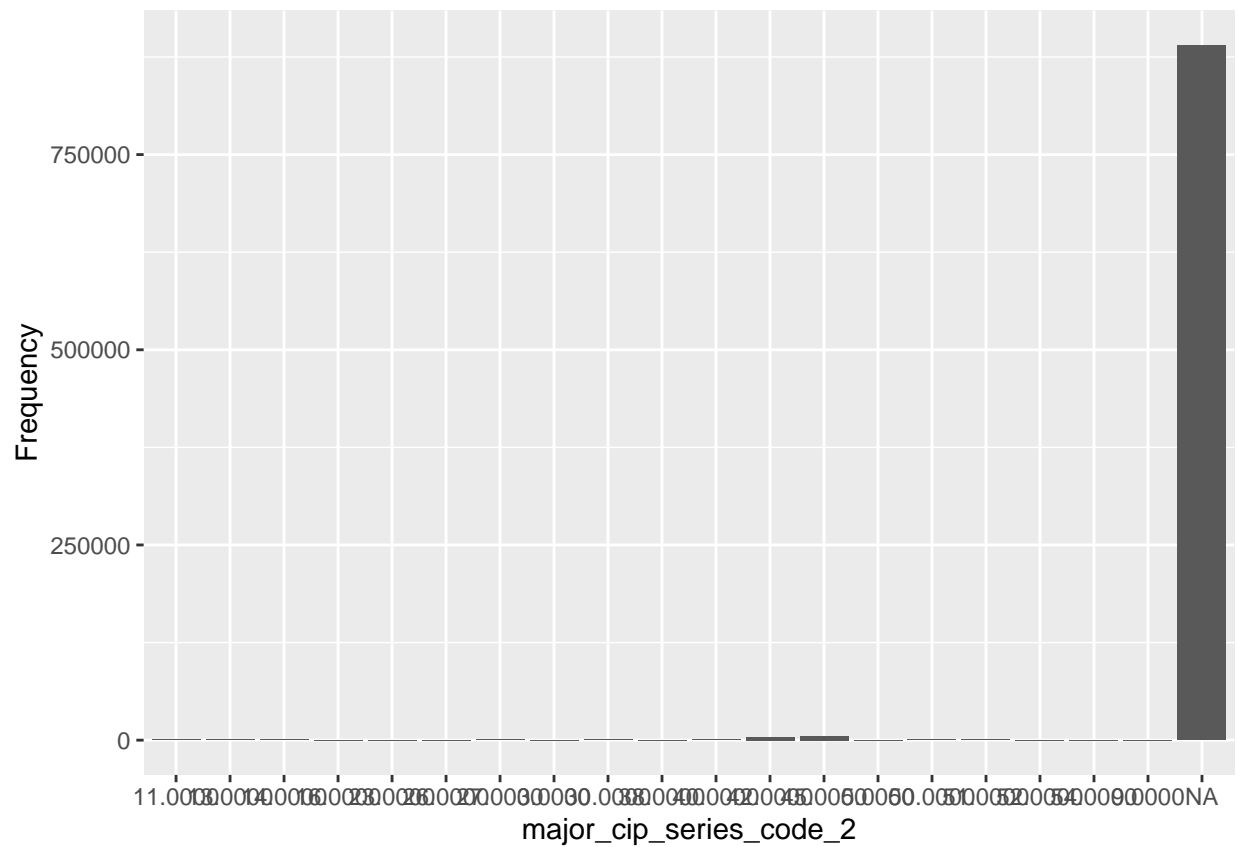
```
[1] -----
[1] Variable: major_cip_code_historical_2, type: character, requested: yes
[1] Values (72 unique): NA, 45.1101, 51.2299, 23.0101, 11.0701, ...
[1] Missing: 98%
[1] Most missing: 2006 100%, Least missing: 2018 93.1%
```



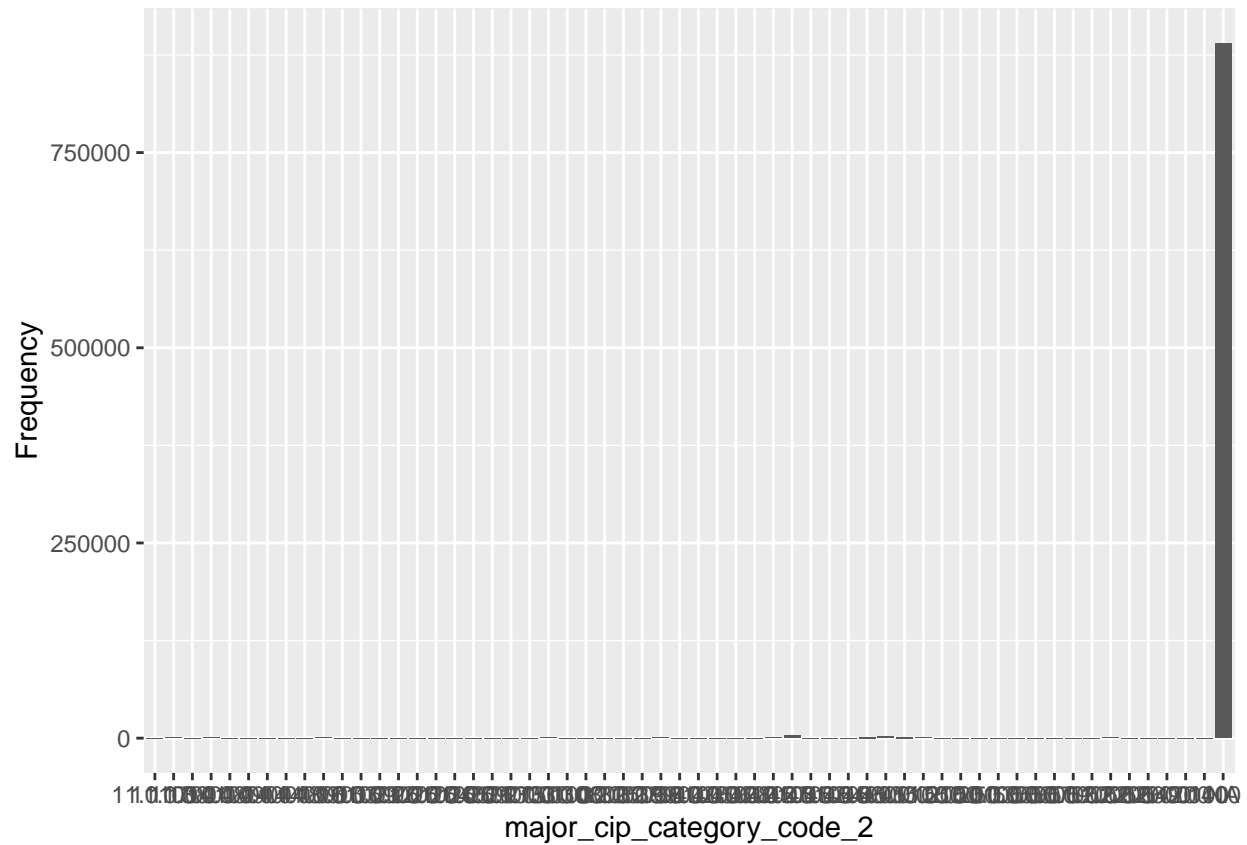
```
[1] -----
[1] Variable: major_cip_name_2, type: character, requested: yes
[1] Values (71 unique): NA, Sociology, Public Health, Other, English Language and Literature, General, (
[1] Missing: 98%
[1] Most missing: 2006 100%, Least missing: 2018 93.1%
```



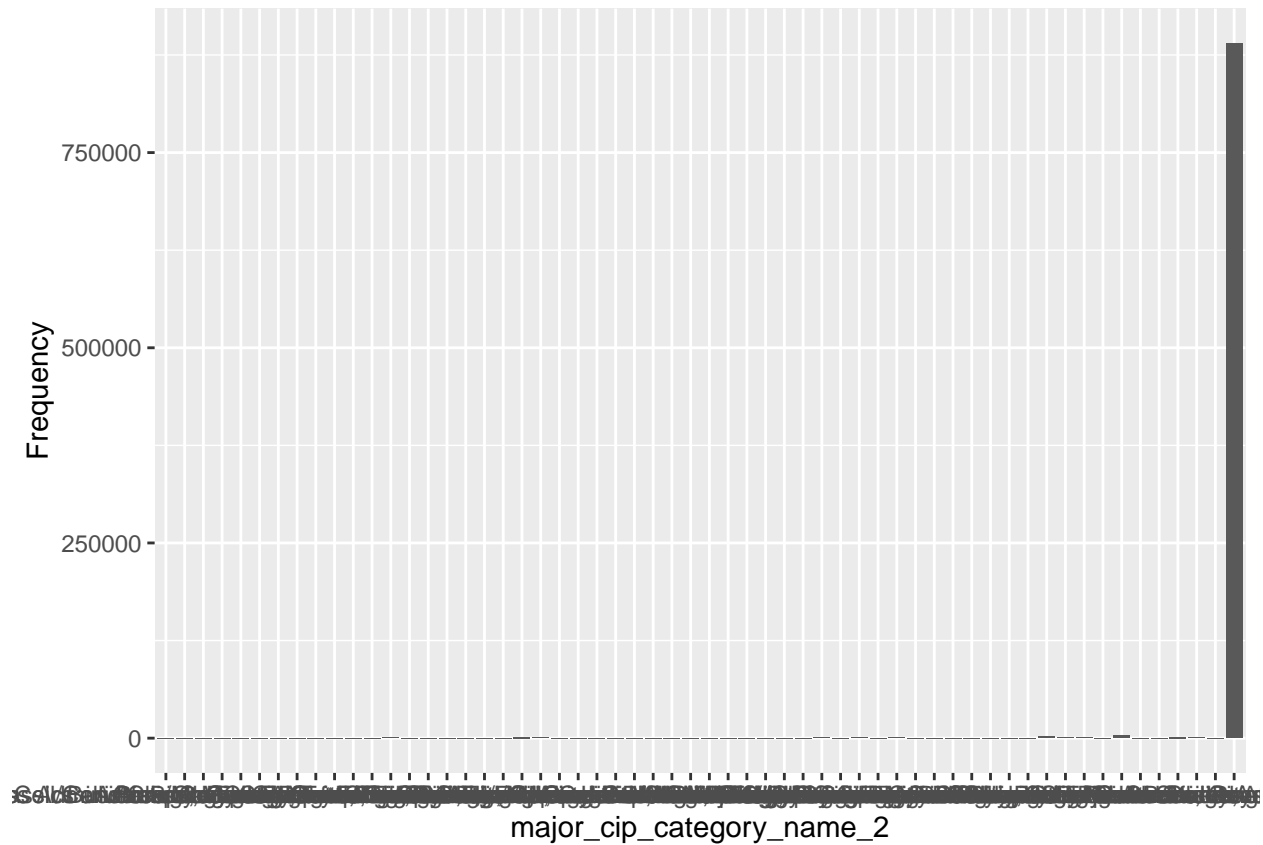
```
[1] -----
[1] Variable: major_cip_series_code_2, type: character, requested: yes
[1] Values (20 unique): NA, 45.0000, 51.0000, 23.0000, 11.0000, ...
[1] Missing: 98%
[1] Most missing: 2006 100%, Least missing: 2018 93.1%
```



```
[1] -----
[1] Variable: major_cip_series_name_2, type: character, requested: yes
[1] Values (21 unique): NA, Social Sciences, Health Professions and Related Programs, English Language &
[1] Missing: 98%
[1] Most missing: 2006 100%, Least missing: 2018 93.1%
```

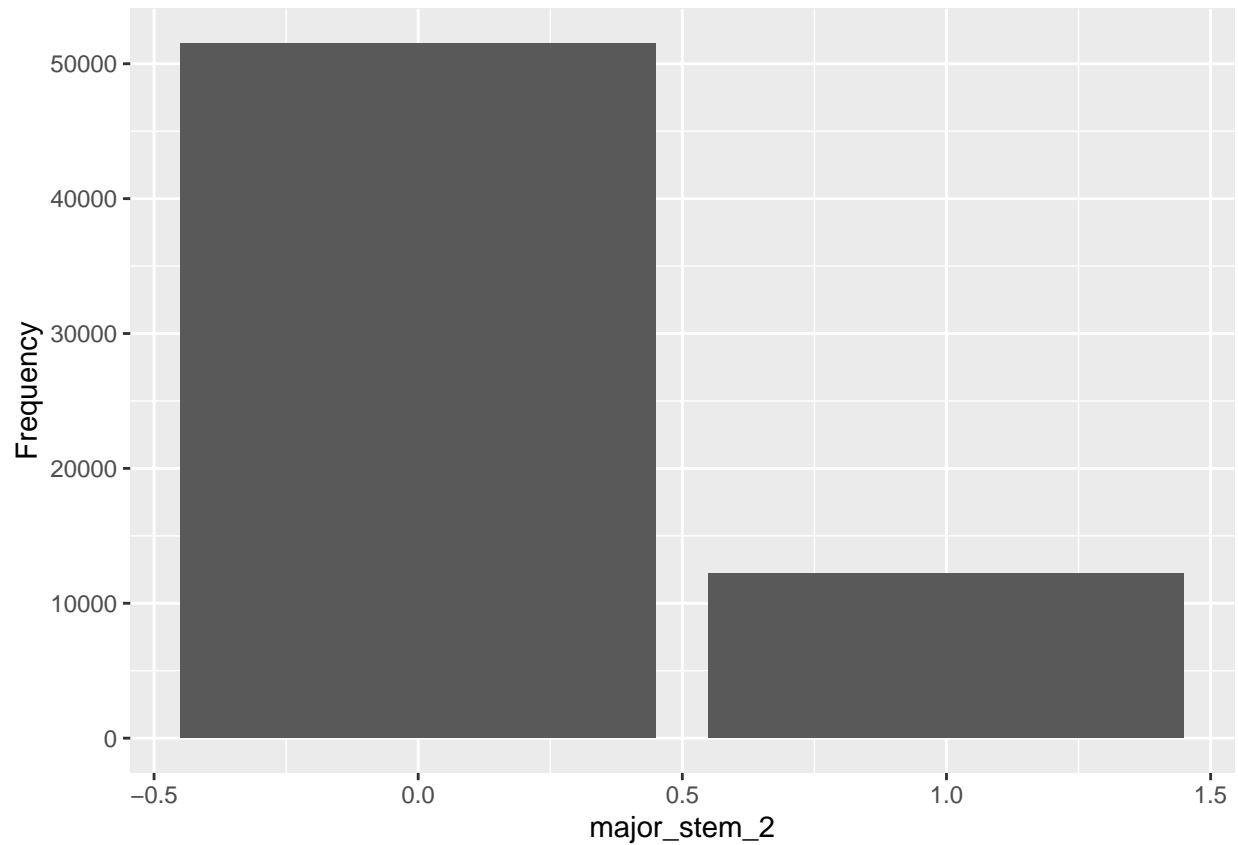



```
[1] -----
[1] Variable: major_cip_category_name_2, type: character, requested: yes
[1] Values (58 unique): NA, Sociology, Public Health, English Language and Literature, General, Computer Science,
[1] Missing: 98%
[1] Most missing: 2006 100%, Least missing: 2018 93.1%
```

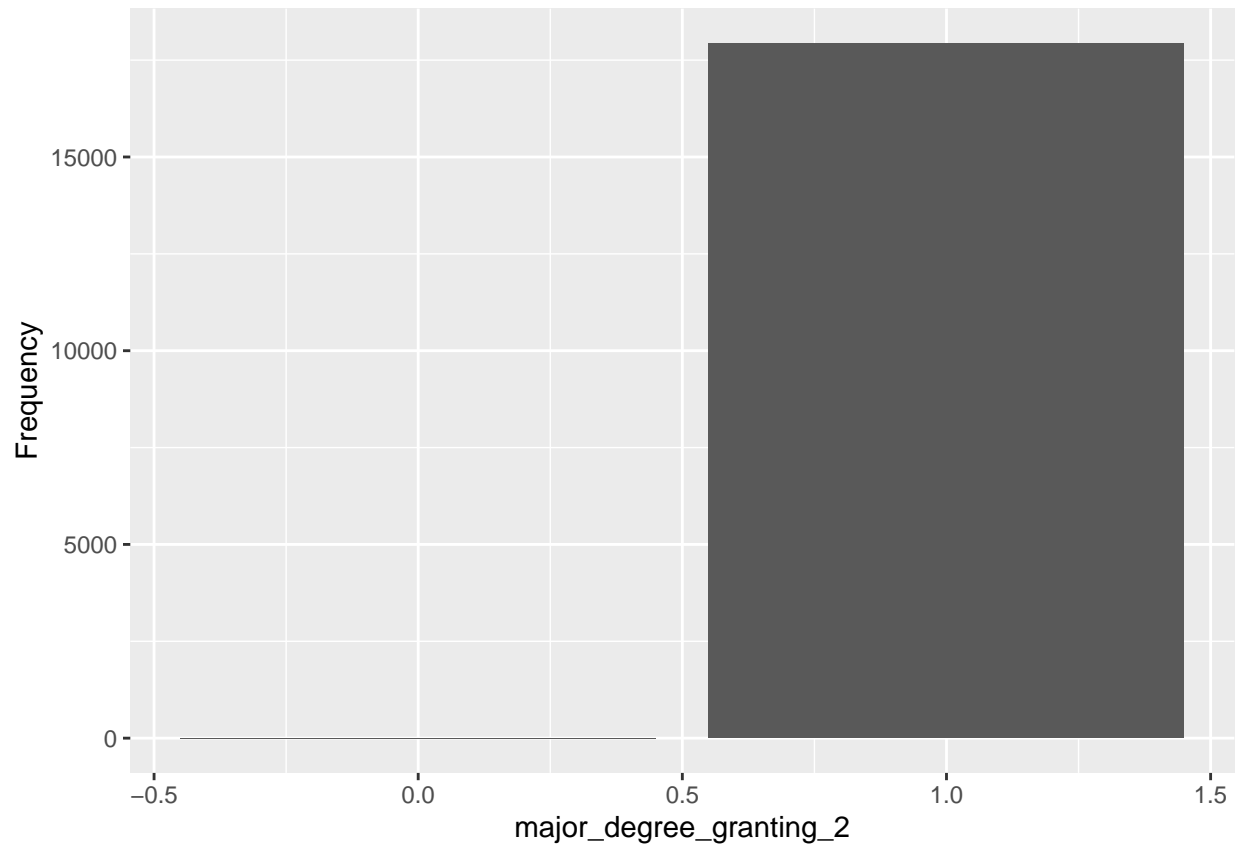
```
[1] -----
[1] Variable: major_stem_2, type: numeric, requested: yes
[1] Values (3 unique): NA, 0, 1
[1] Missing: 93%
[1] Most missing: 2006 95.9%, Least missing: 2015 92.3%
```

Warning: Removed 844109 rows containing non-finite values ('stat_count()').

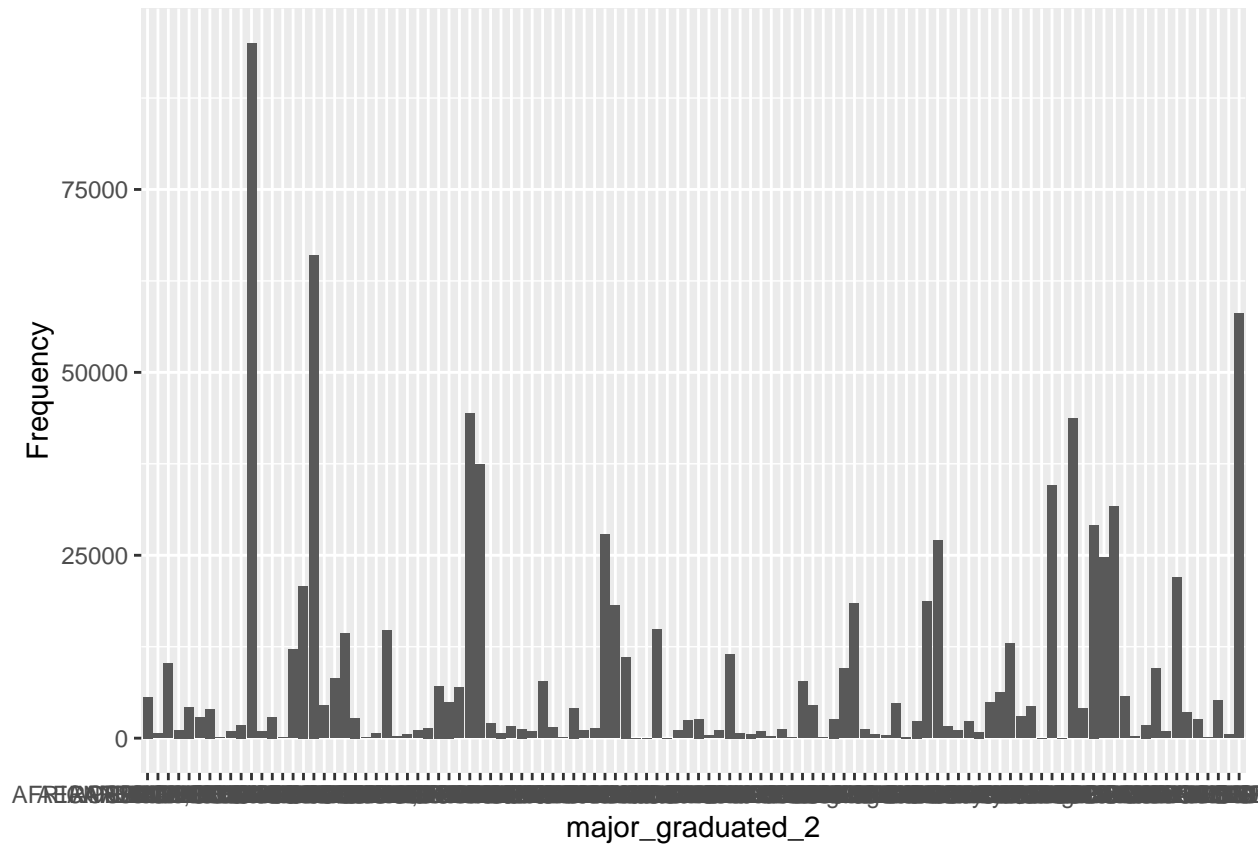


```
[1] -----  
[1] Variable: major_degree_granting_2, type: numeric, requested: yes  
[1] Values (3 unique): NA, 1, 0  
[1] Missing: 98%  
[1] Most missing: 2006 100%, Least missing: 2018 93.1%
```

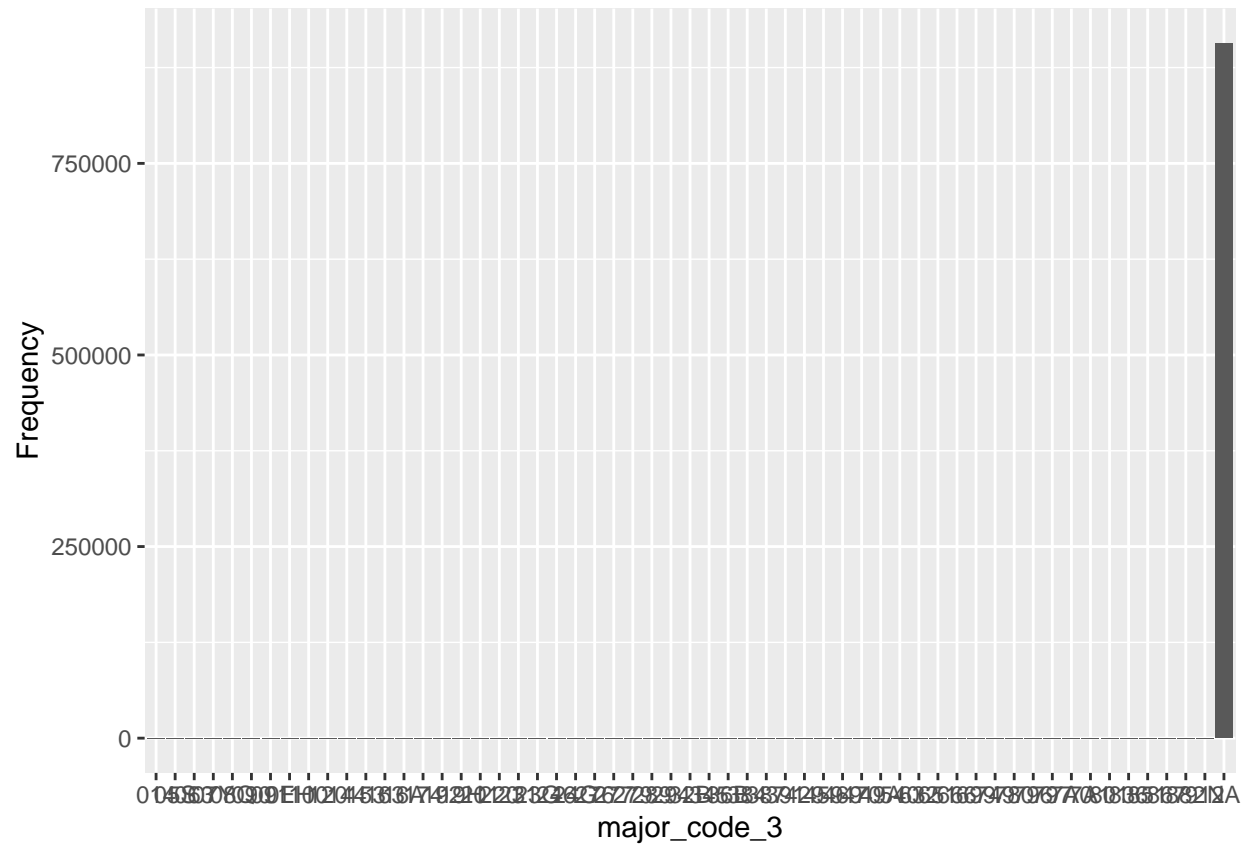
Warning: Removed 889895 rows containing non-finite values ('stat_count()').



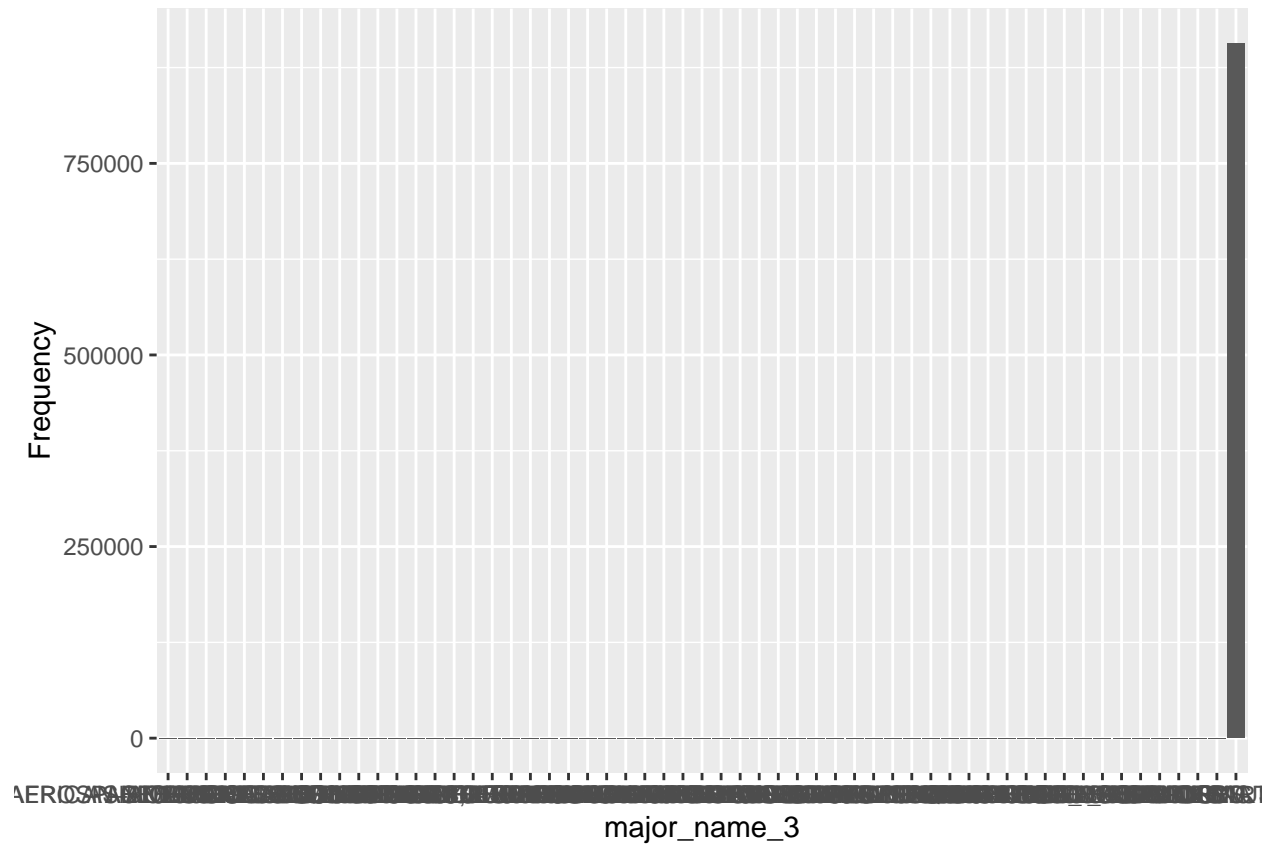
```
[1] -----  
[1] Variable: major_graduated_2, type: character, requested: yes  
[1] Values (106 unique): NURSING SCIENCE, COMPUTER SCIENCE, CHICANO/LATINO STD, BUSINESS ECONOMICS, NA,  
[1] Missing: 6.4%  
[1] Most missing: 2018 13.9%, Least missing: 2006 2.8%
```



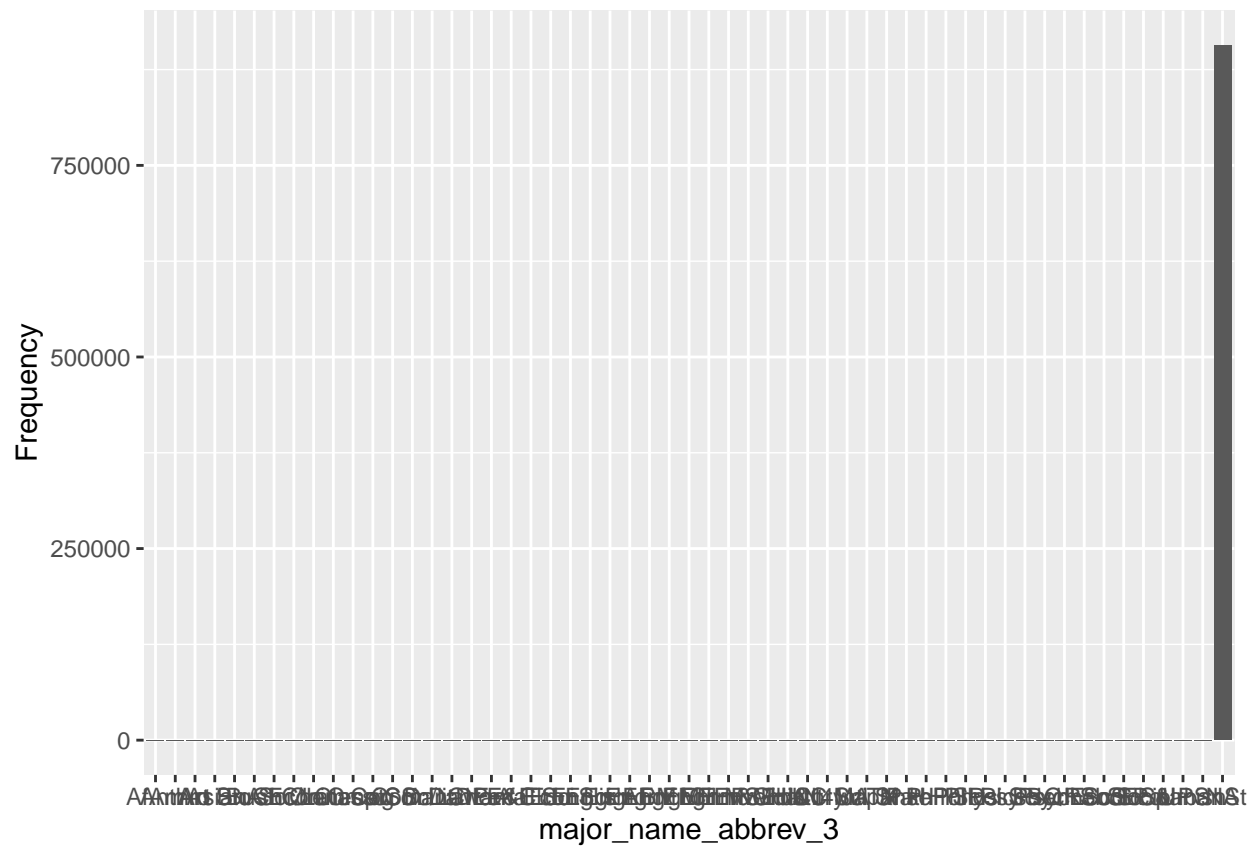
```
[1] -----
[1] Variable: major_code_3, type: character, requested: yes
[1] Values (57 unique): NA, 277, 632, 851, 867, ...
[1] Missing: 99.9%
```



```
[1] -----
[1] Variable: major_name_3, type: character, requested: yes
[1] Values (57 unique): NA, MECHANICAL ENGINEERING, PSYCHOLOGY AND SOCIAL BEHAVIOR, SOCIAL ECOLOGY, SOC
[1] Missing: 99.9%
```



```
[1] -----
[1] Variable: major_name_abbrev_3, type: character, requested: yes
[1] Values (55 unique): NA, Engr ME, Psy Beh, SocEcol, Sociol, ...
[1] Missing: 99.9%
```

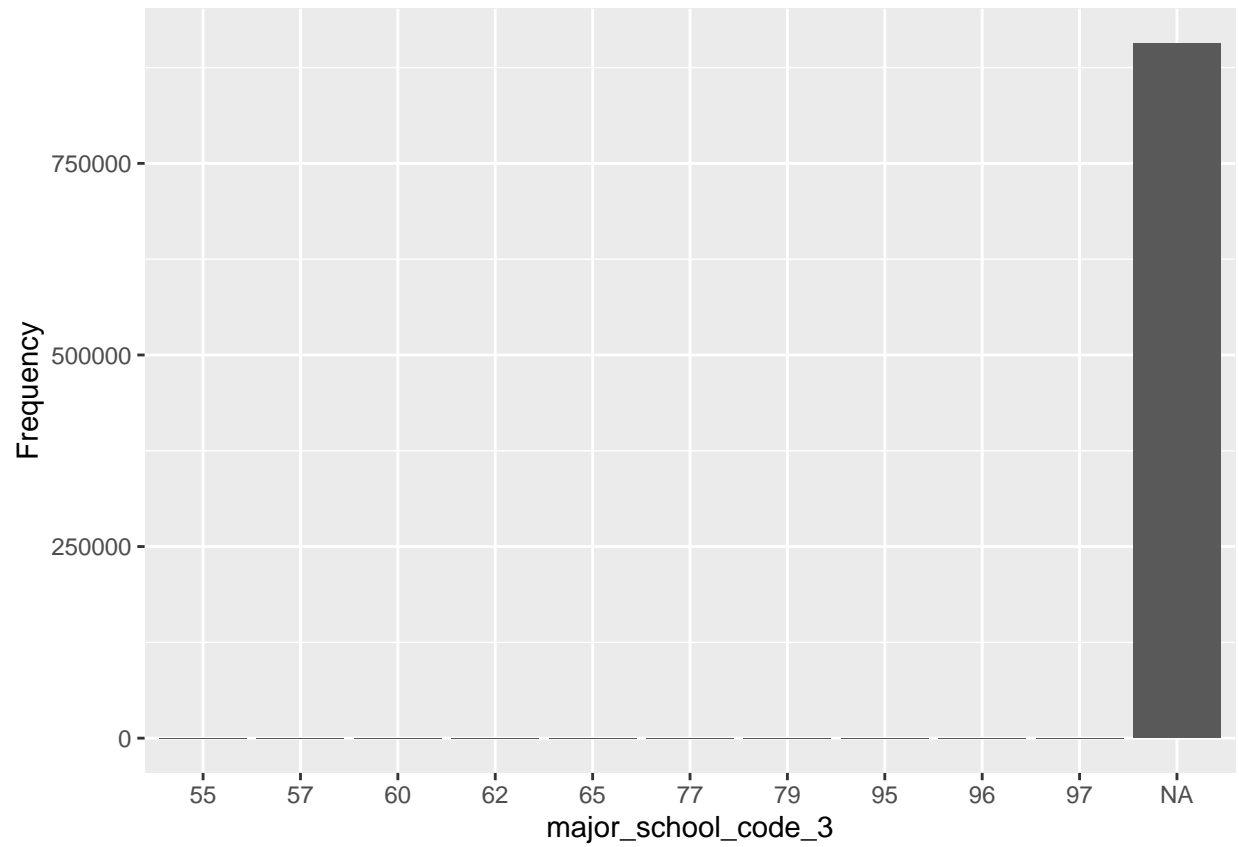


```
[1] -----
[1] Variable: major_school_code_3, type: character, requested: yes
[1] Values (11 unique): NA, 77, 97, 65, 60, ...
[1] Missing: 99.9%
```

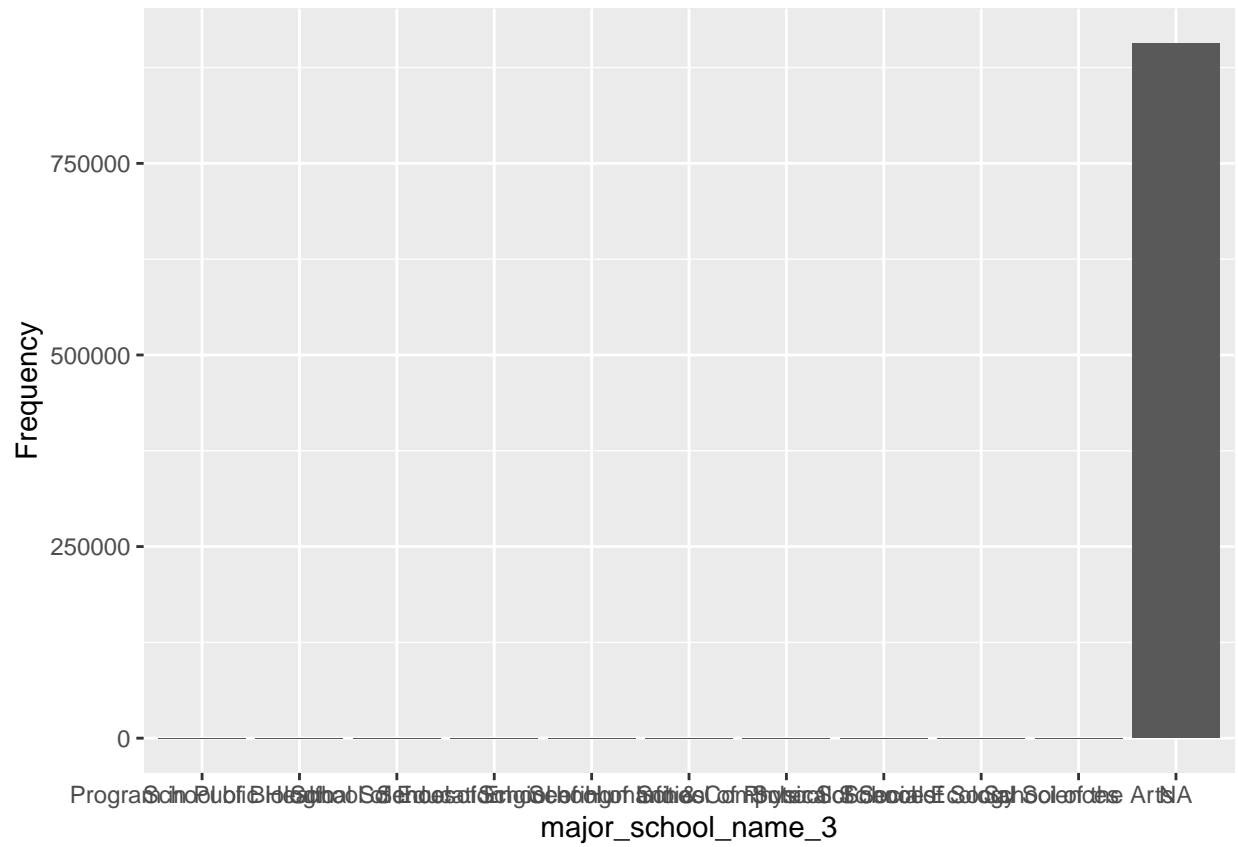
```
[1] Variable: major_school_code_3, type: character, requested: yes
```

```
[1] Values (11 unique): NA, 77, 97, 65, 60, ...
```

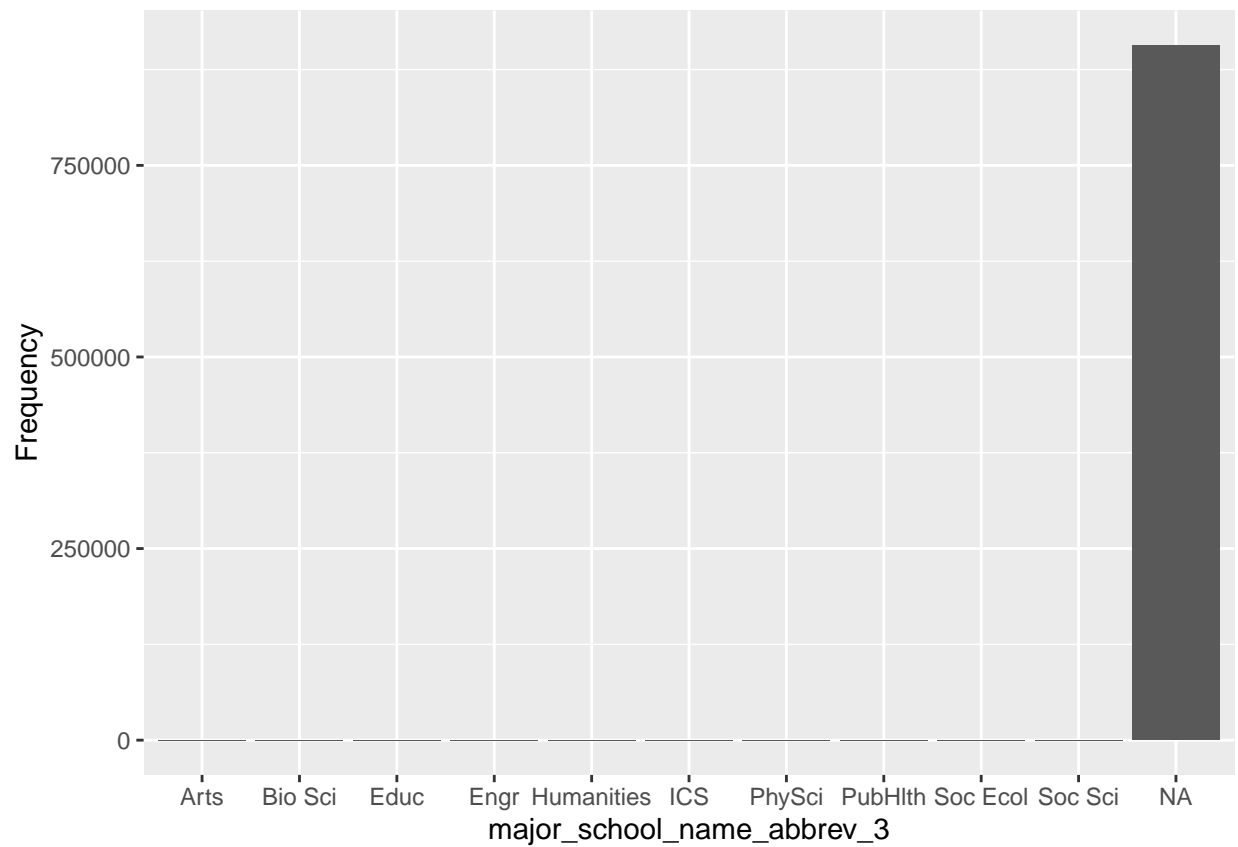
```
[1] Missing: 99.9%
```



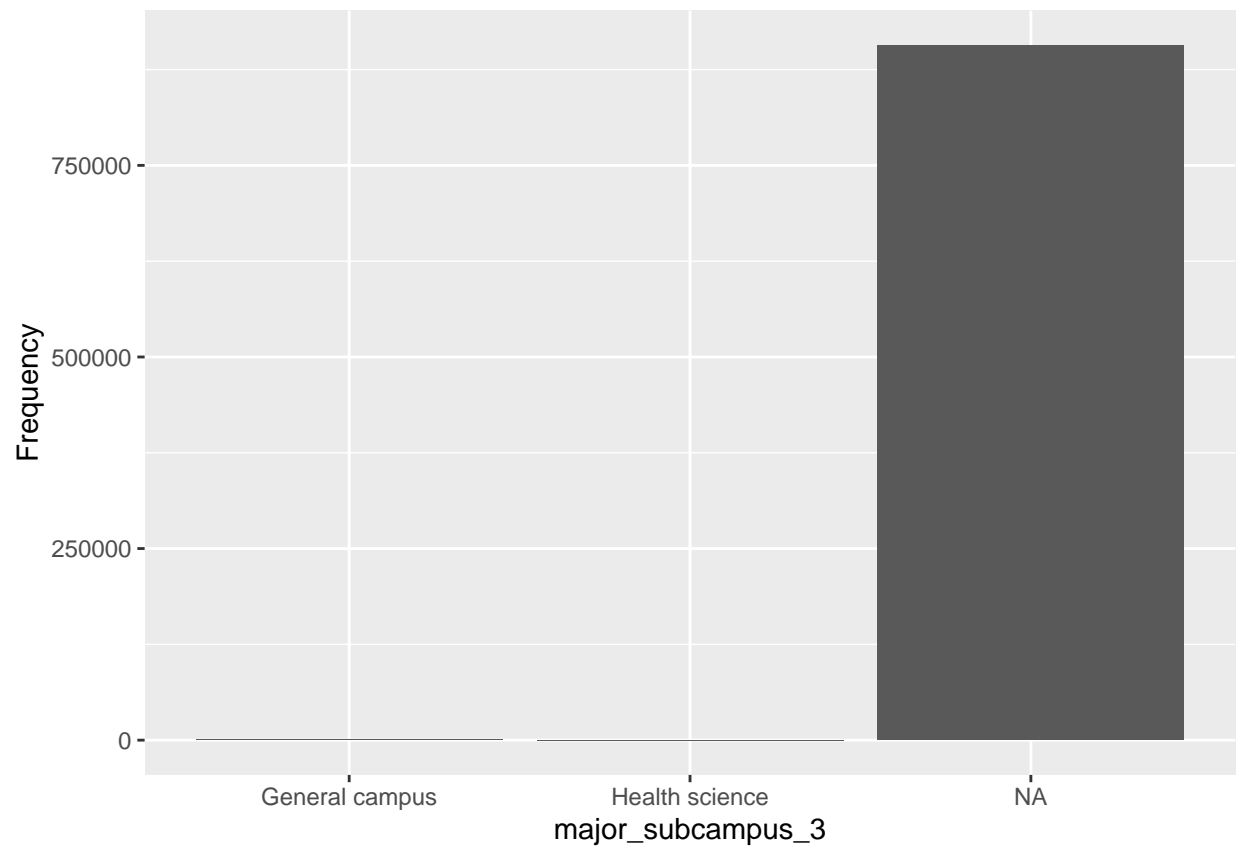
```
[1] -----  
[1] Variable: major_school_name_3, type: character, requested: yes  
[1] Values (11 unique): NA, School of Engineering, School of Social Ecology, School of Social Sciences,  
[1] Missing: 99.9%
```

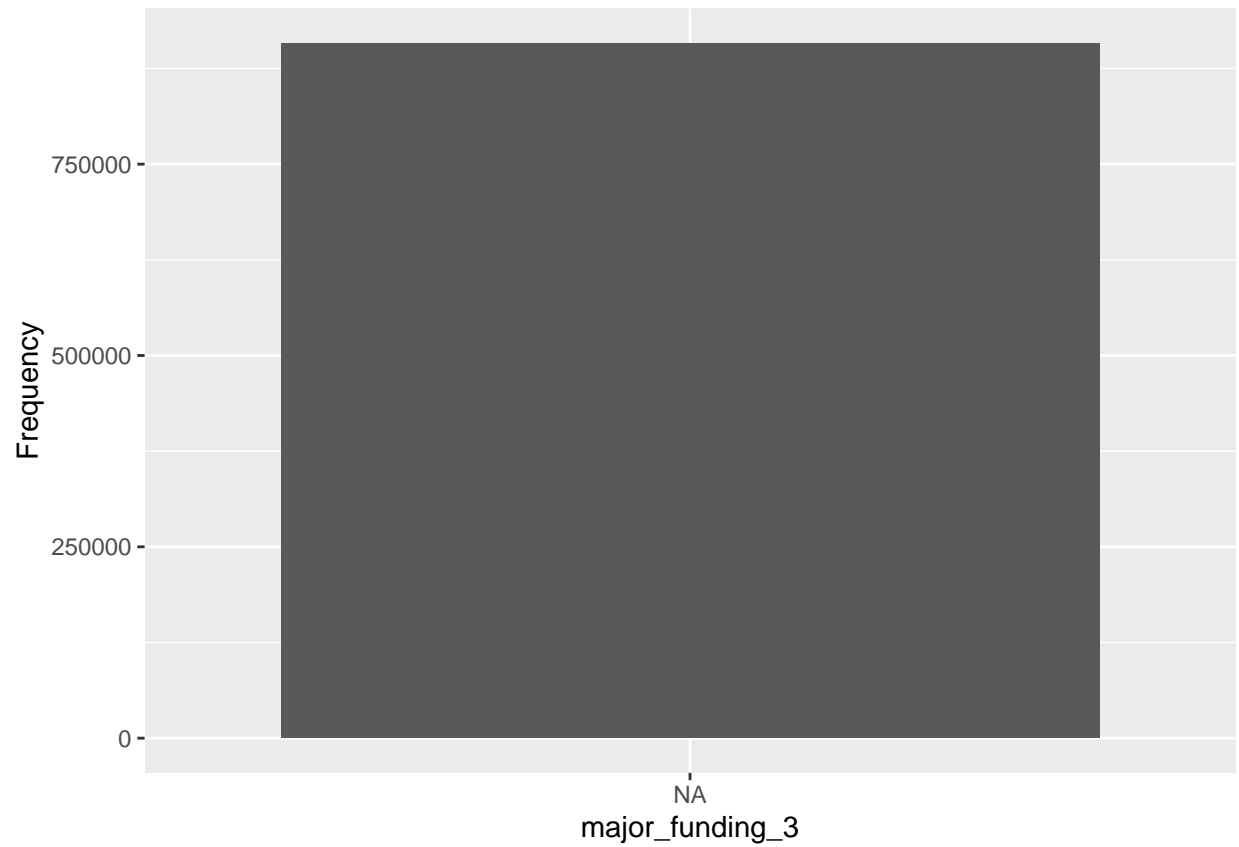
```
[1] -----
[1] Variable: major_school_name_abbrev_3, type: character, requested: yes
[1] Values (11 unique): NA, Engr, Soc Ecol, Soc Sci, Humanities, ...
[1] Missing: 99.9%
```



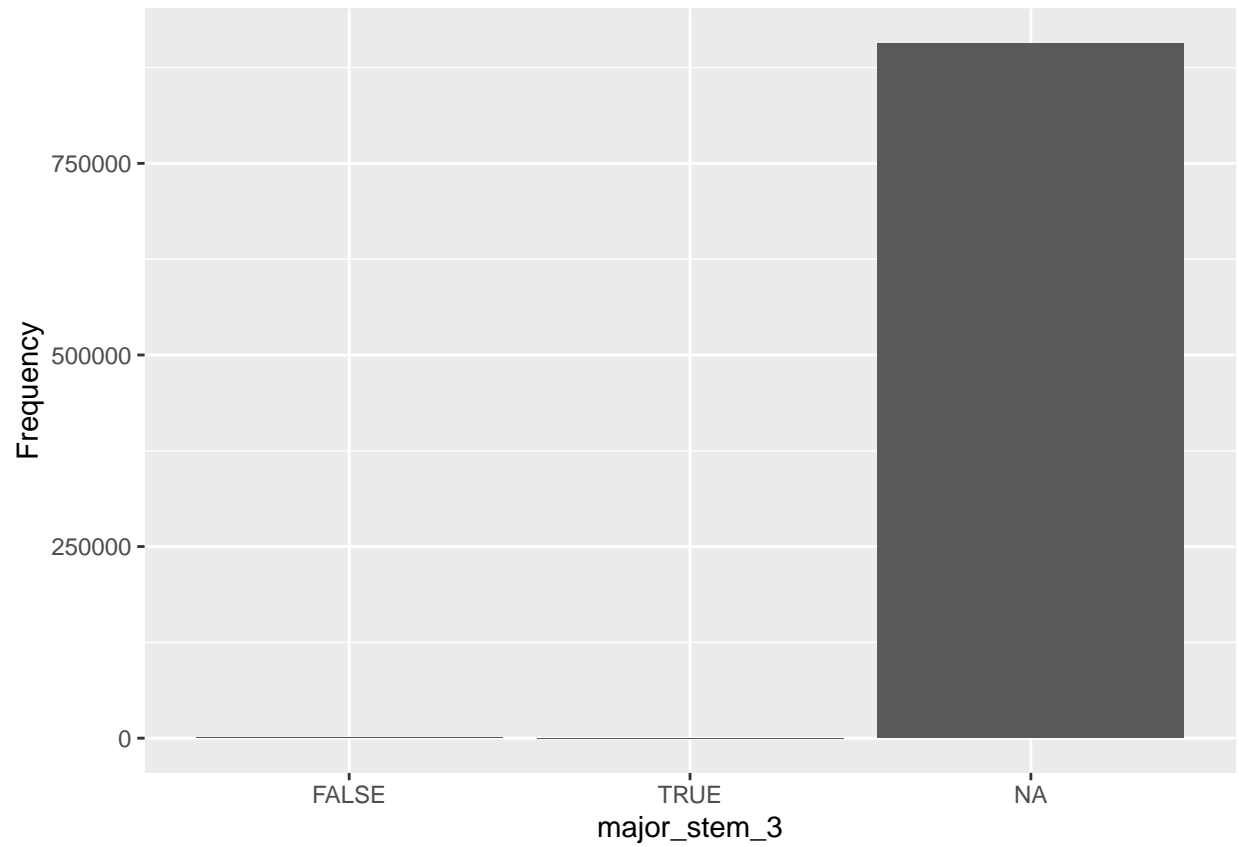
```
[1] -----
[1] Variable: major_subcampus_3, type: character, requested: yes
[1] Values (3 unique): NA, General campus, Health science
[1] Missing: 99.9%
```



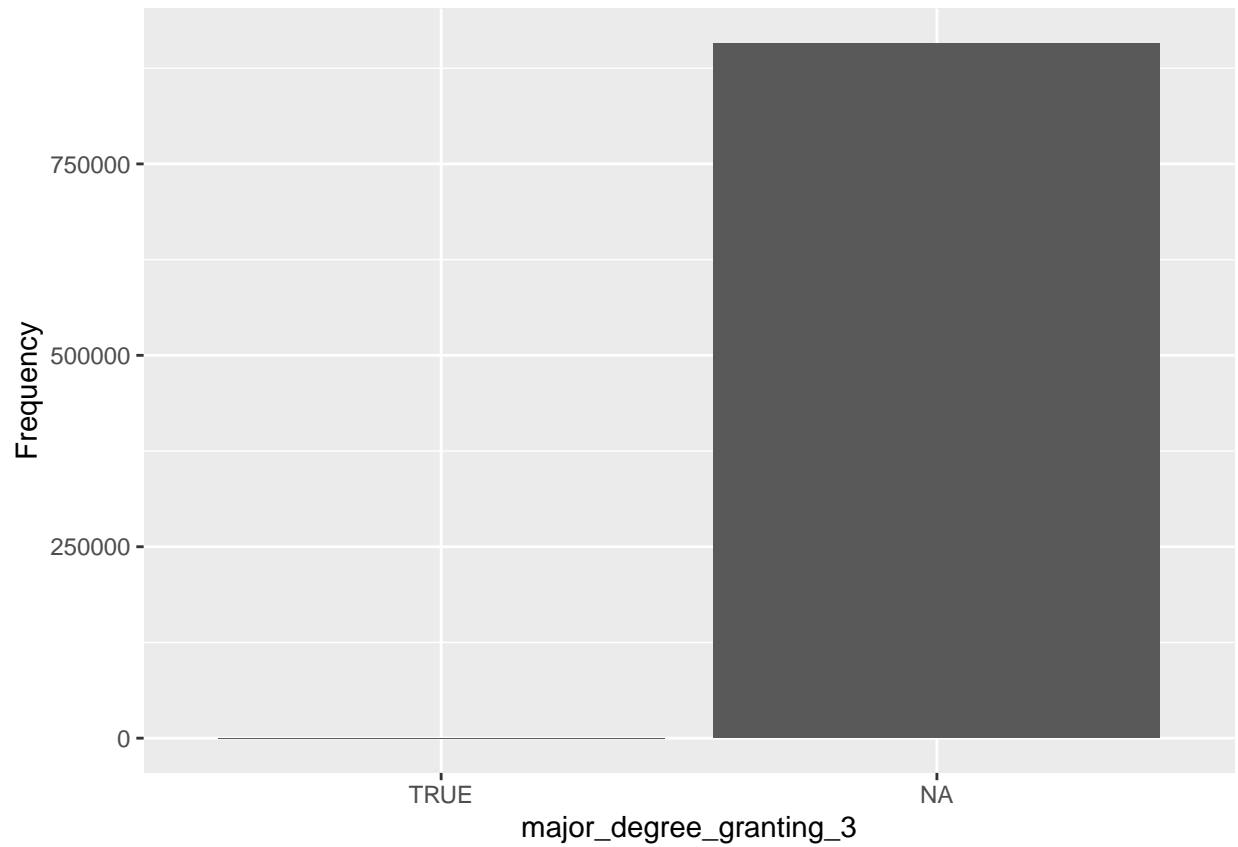
```
[1] -----  
[1] Variable: major_funding_3, type: logical, requested: yes  
[1] Values (1 unique): NA  
[1] Missing: 100%
```



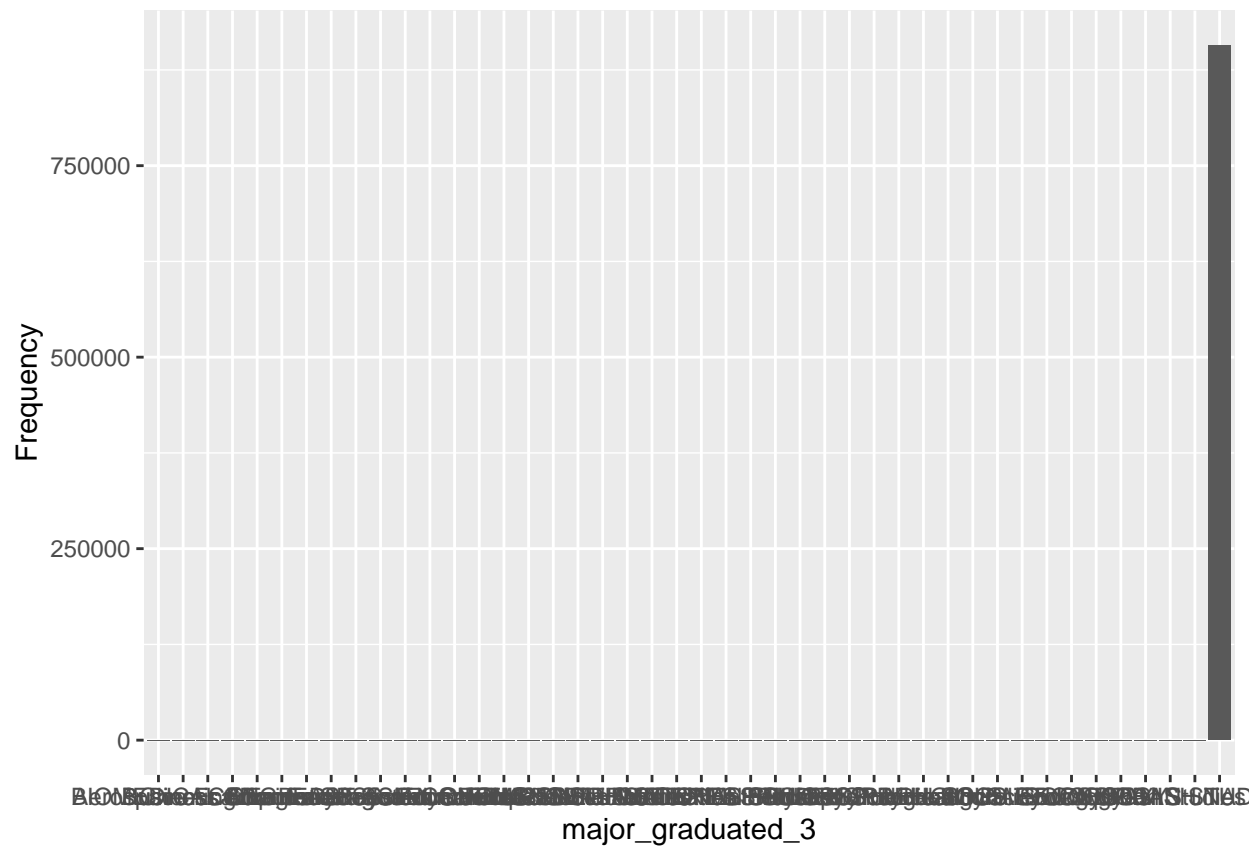
```
[1] -----  
[1] Variable: major_stem_3, type: logical, requested: yes  
[1] Values (3 unique): NA, TRUE, FALSE  
[1] Missing: 99.9%
```



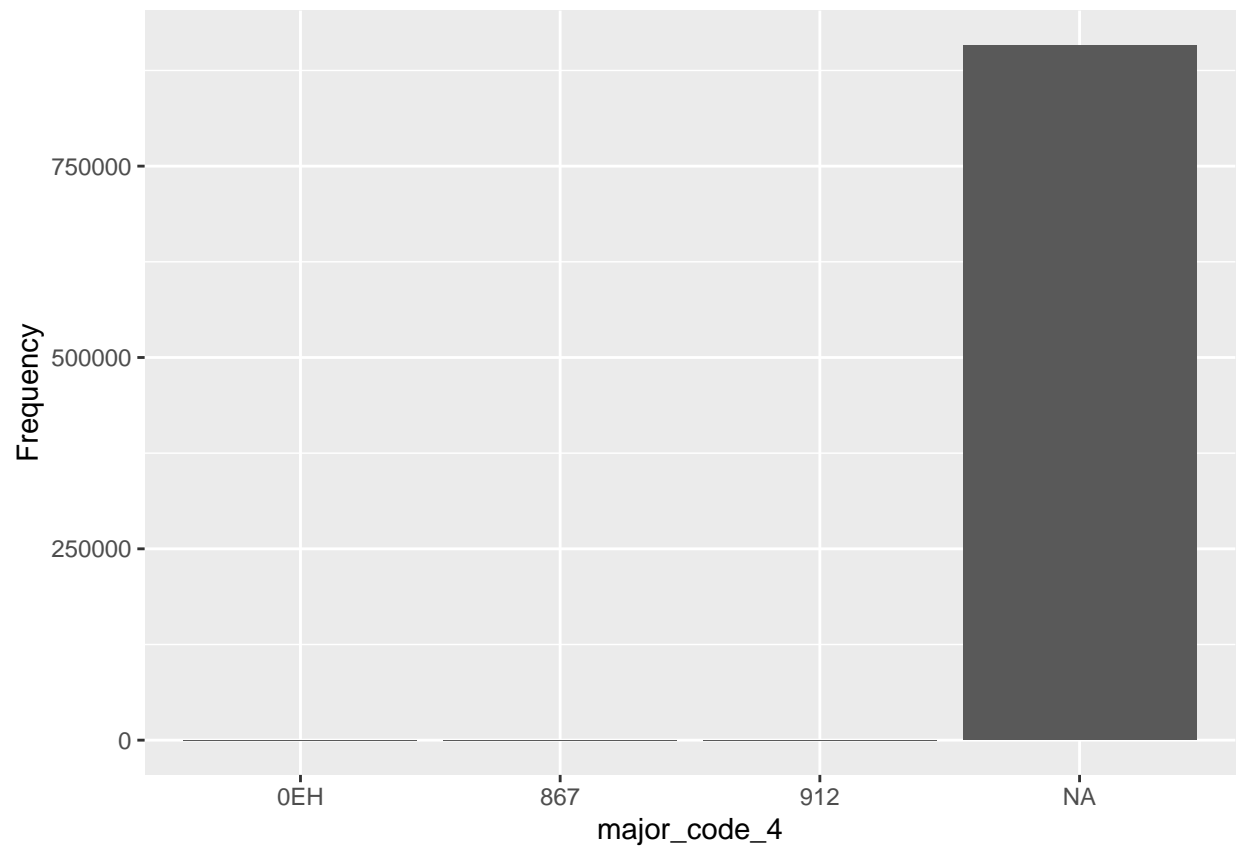
```
[1] -----  
[1] Variable: major_degree_granting_3, type: logical, requested: yes  
[1] Values (2 unique): NA, TRUE  
[1] Missing: 100%
```



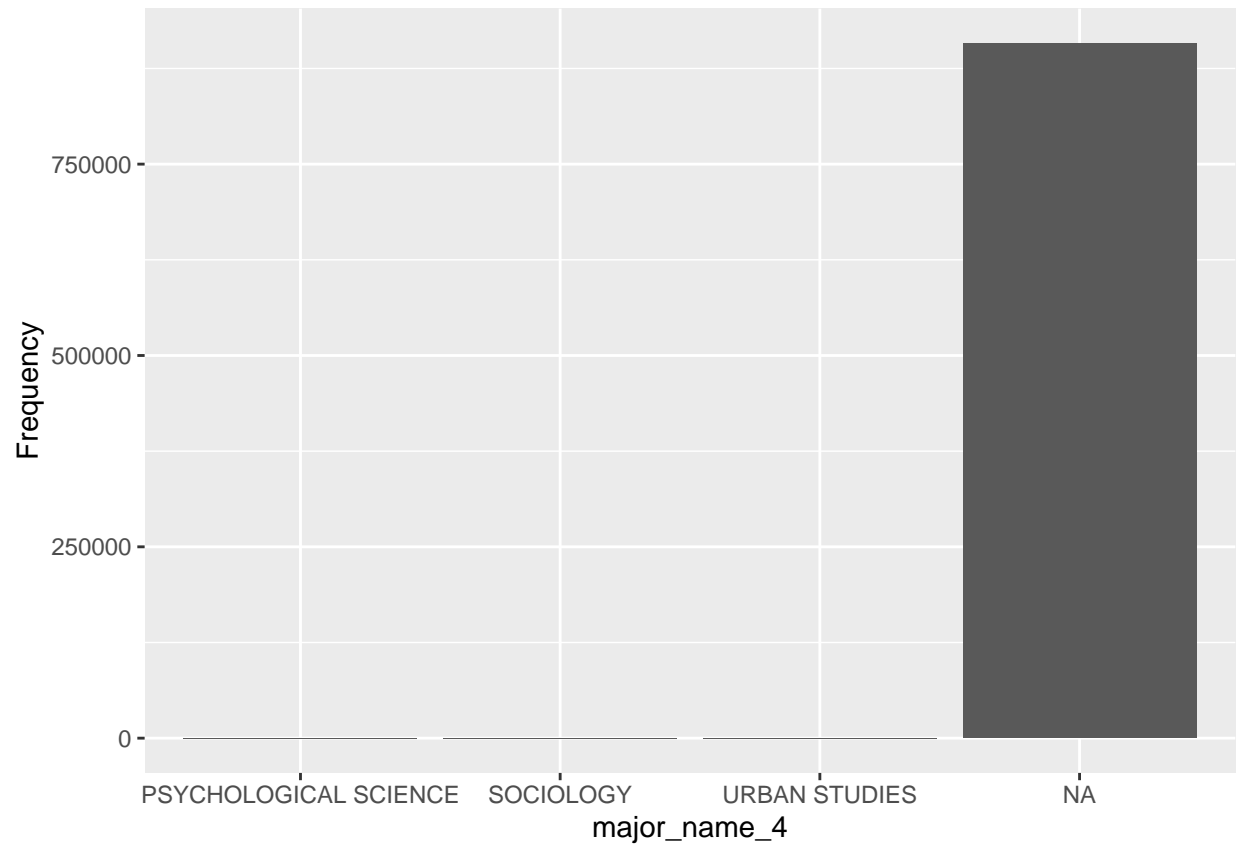
```
[1] -----  
[1] Variable: major_graduated_3, type: character, requested: yes  
[1] Values (44 unique): NA, Aerospace Engr, Crim, Law & Society, Psych & Soc Behavior, Psychological Sc  
[1] Missing: 99.9%
```



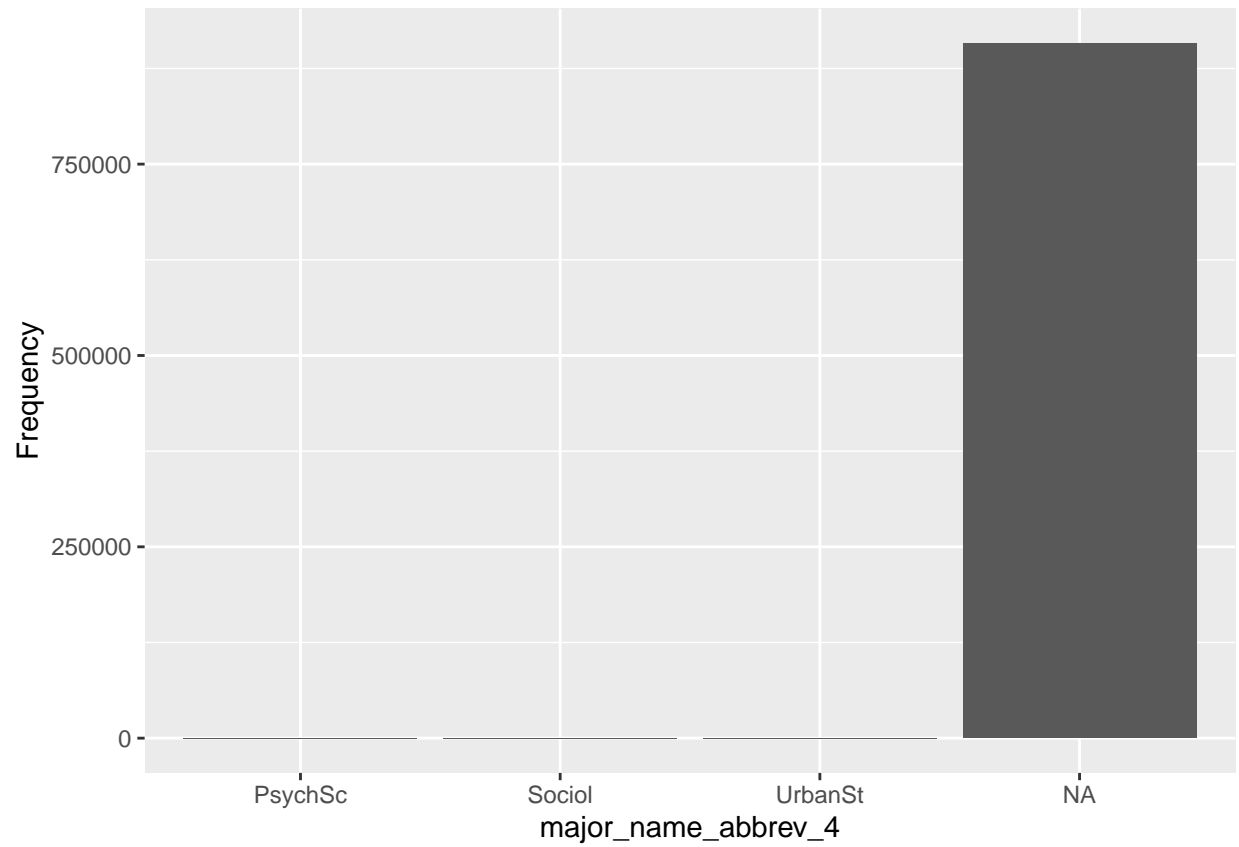
```
[1] -----
[1] Variable: major_code_4, type: character, requested: yes
[1] Values (4 unique): NA, 912, 0EH, 867
[1] Missing: 100%
```



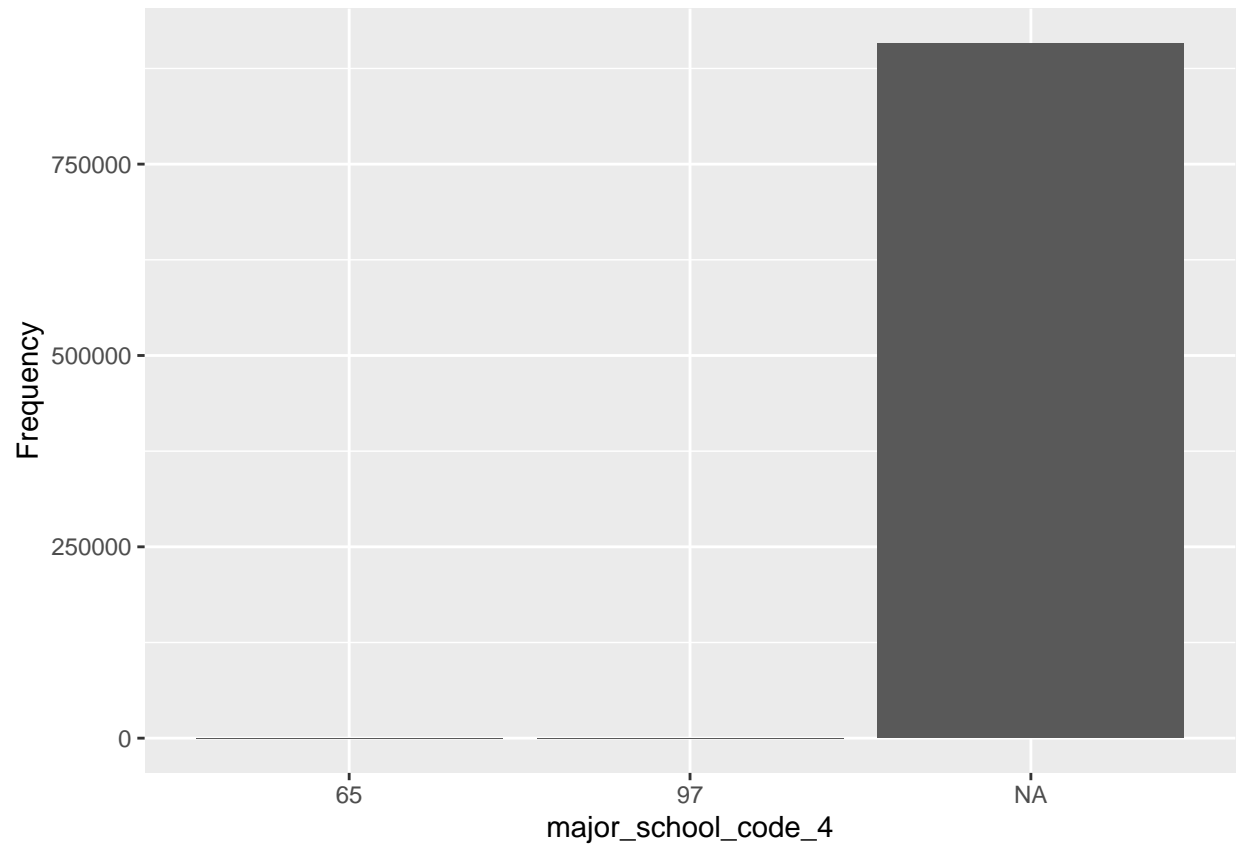
```
[1] -----  
[1] Variable: major_name_4, type: character, requested: yes  
[1] Values (4 unique): NA, URBAN STUDIES, PSYCHOLOGICAL SCIENCE, SOCIOLOGY  
[1] Missing: 100%
```

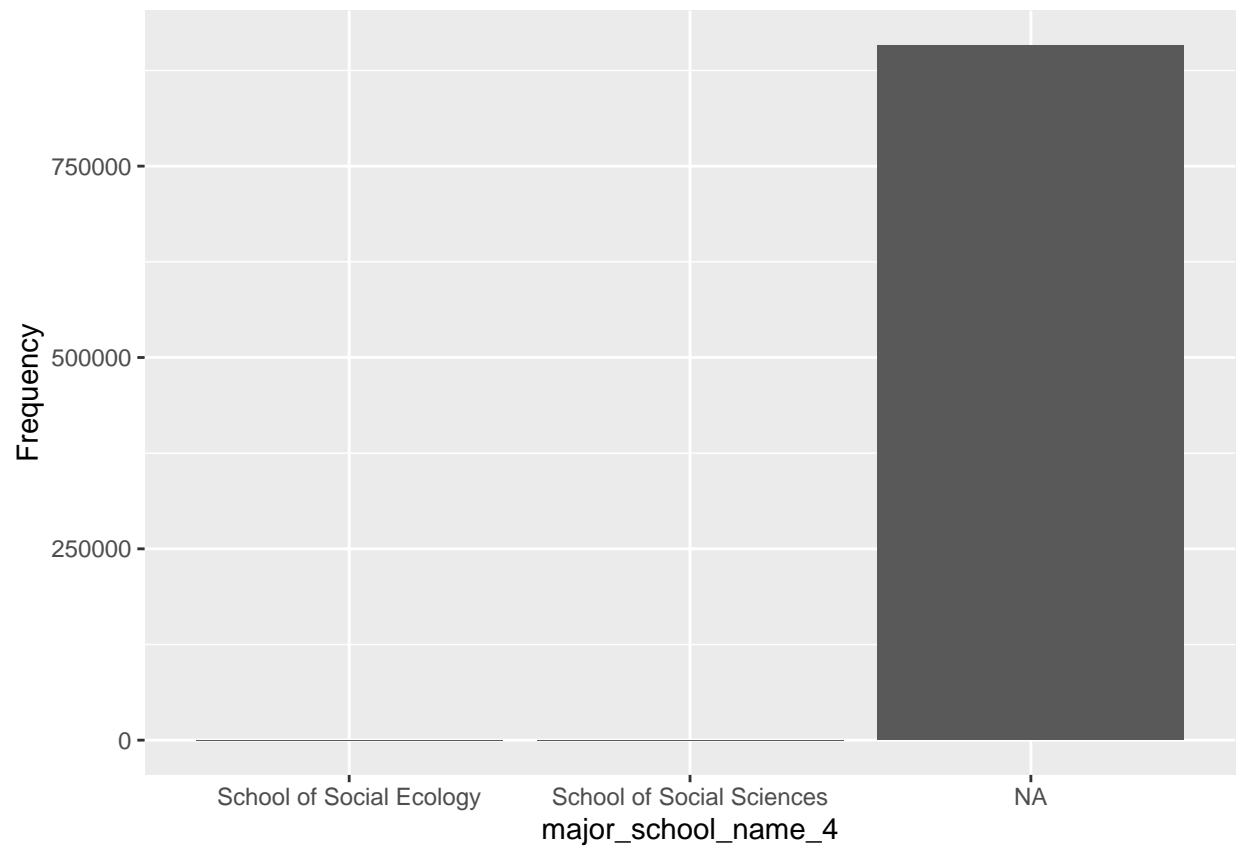
```
[1] -----  
[1] Variable: major_name_abbrev_4, type: character, requested: yes  
[1] Values (4 unique): NA, UrbanSt, PsychSc, Sociol  
[1] Missing: 100%
```



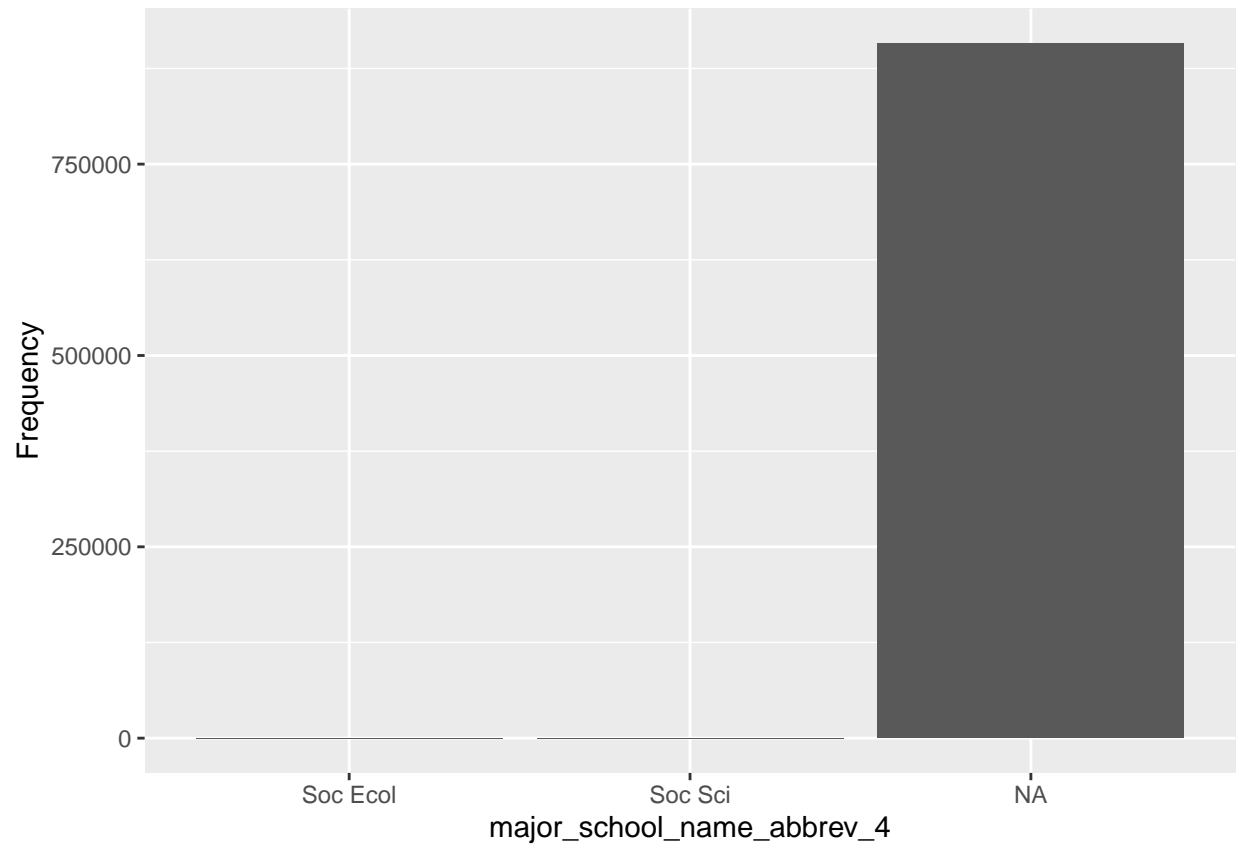
```
[1] -----  
[1] Variable: major_school_code_4, type: character, requested: yes  
[1] Values (3 unique): NA, 97, 65  
[1] Missing: 100%
```



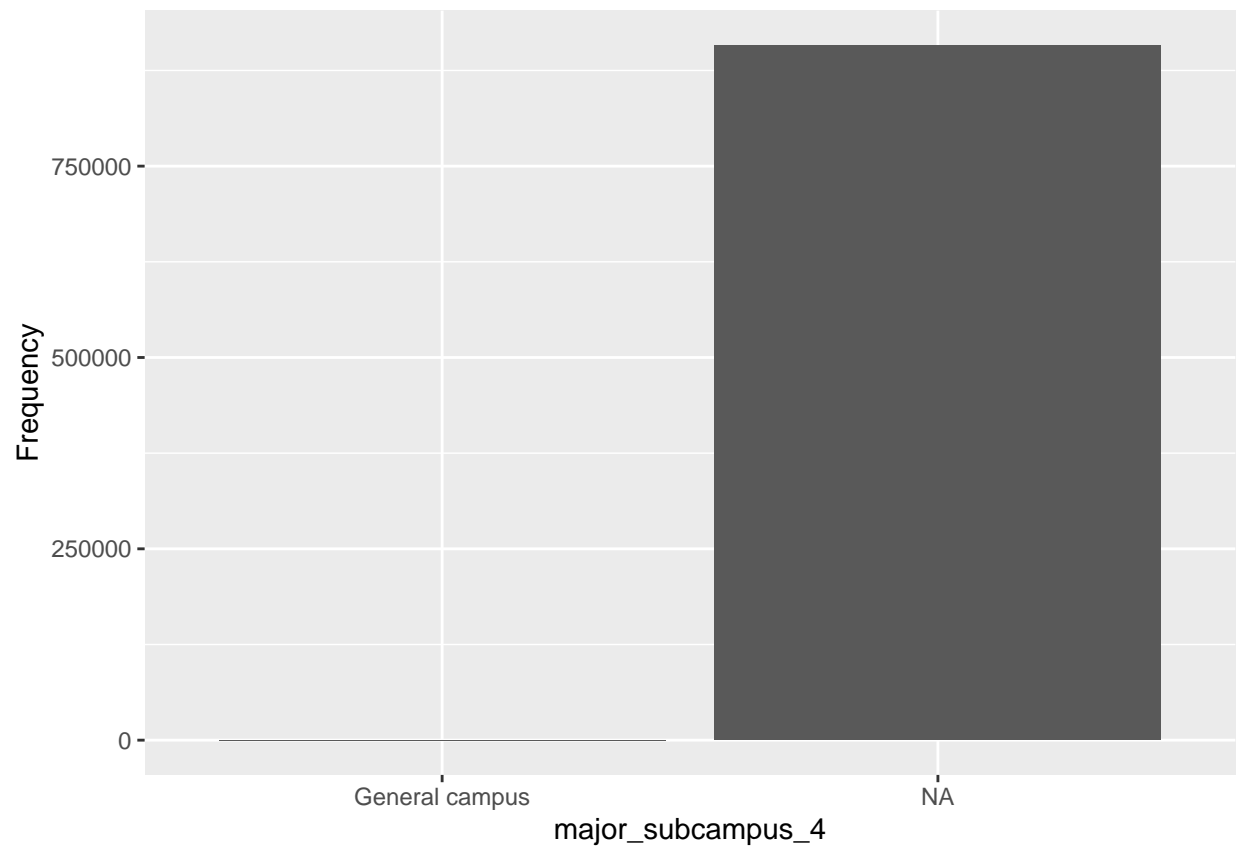
```
[1] -----  
[1] Variable: major_school_name_4, type: character, requested: yes  
[1] Values (3 unique): NA, School of Social Ecology, School of Social Sciences  
[1] Missing: 100%
```



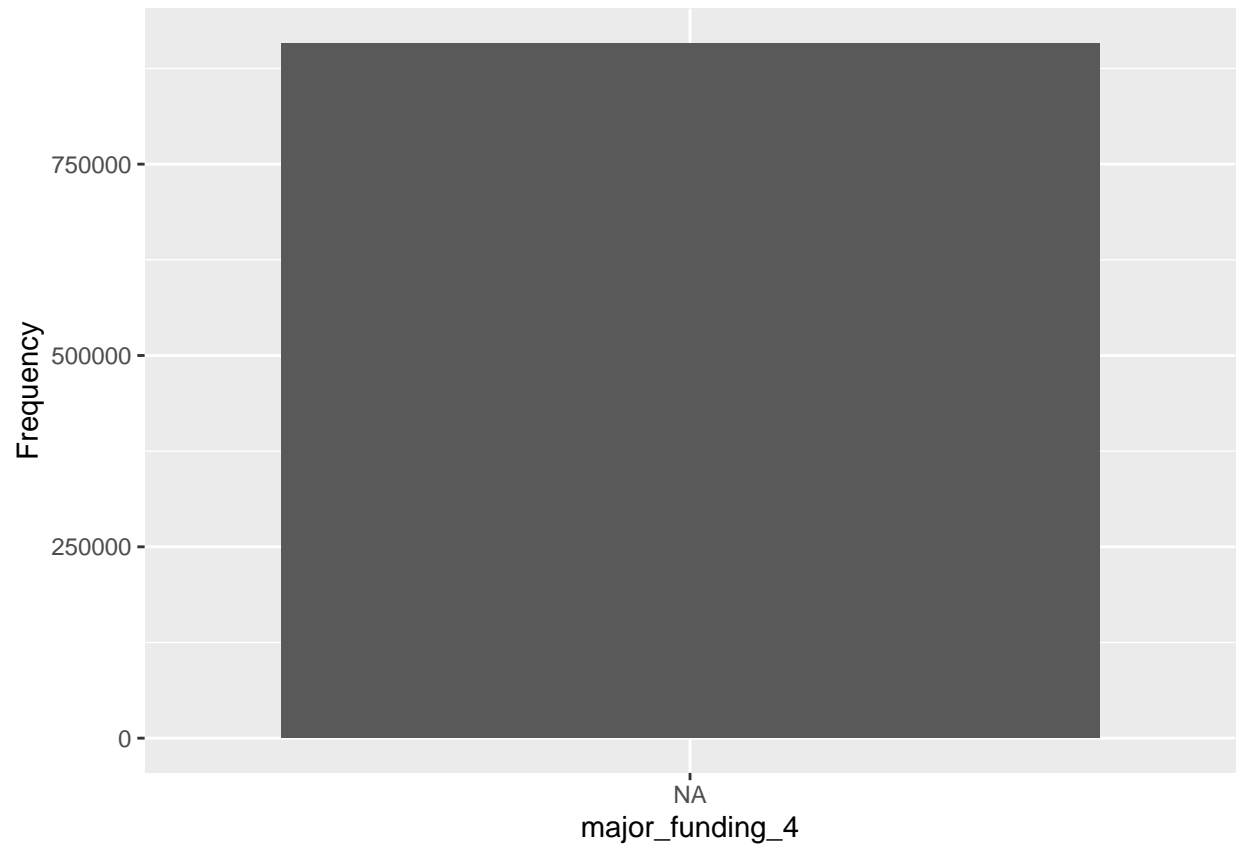
```
[1] -----  
[1] Variable: major_school_name_abbrev_4, type: character, requested: yes  
[1] Values (3 unique): NA, Soc Ecol, Soc Sci  
[1] Missing: 100%
```



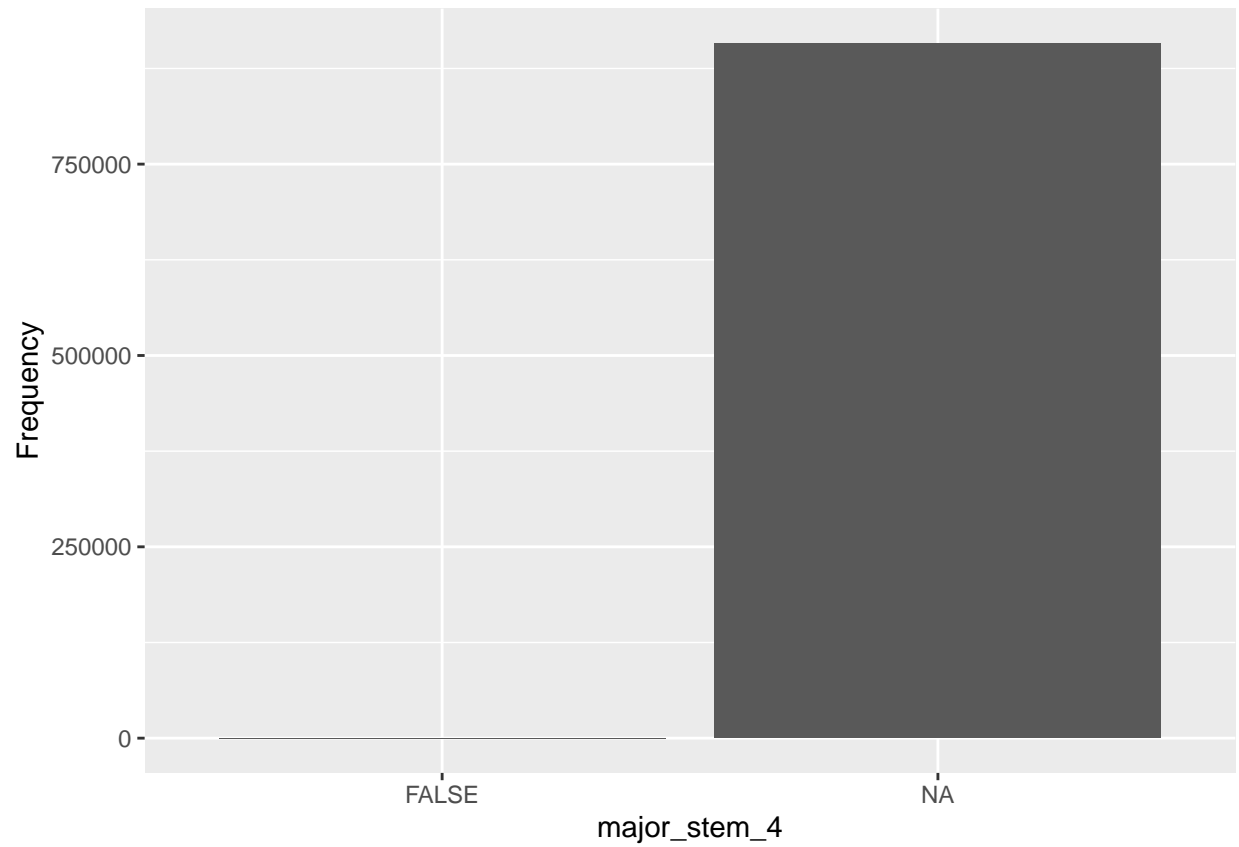
```
[1] -----  
[1] Variable: major_subcampus_4, type: character, requested: yes  
[1] Values (2 unique): NA, General campus  
[1] Missing: 100%
```



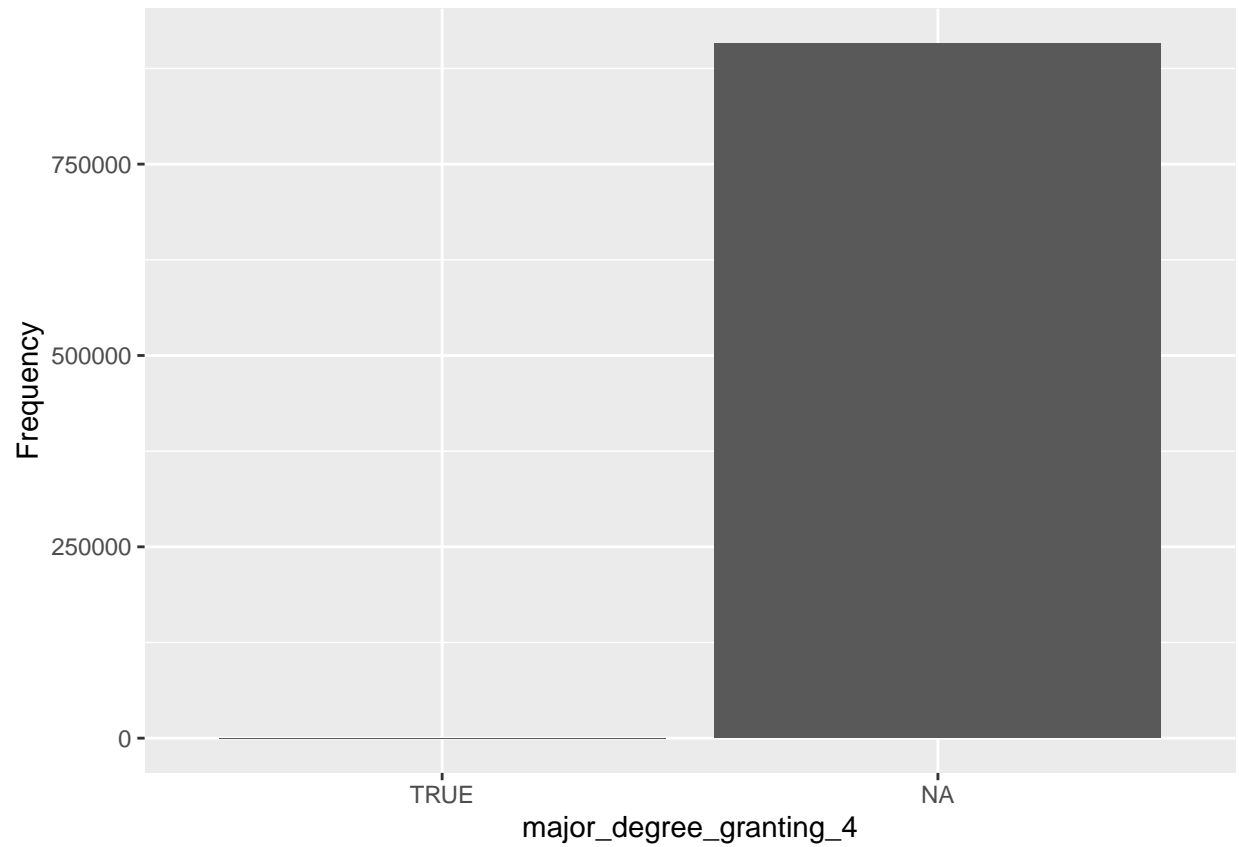
```
[1] -----  
[1] Variable: major_funding_4, type: logical, requested: yes  
[1] Values (1 unique): NA  
[1] Missing: 100%
```



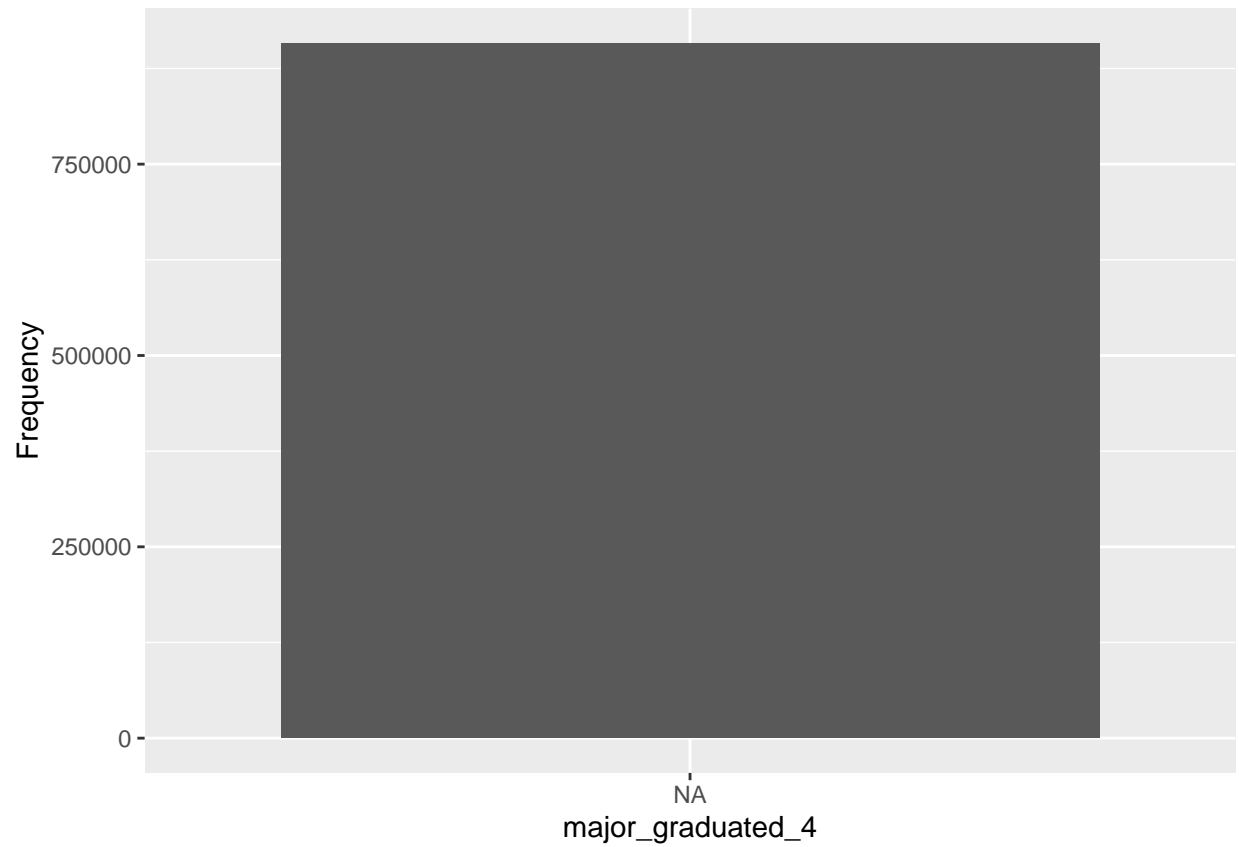
```
[1] -----  
[1] Variable: major_stem_4, type: logical, requested: yes  
[1] Values (2 unique): NA, FALSE  
[1] Missing: 100%
```



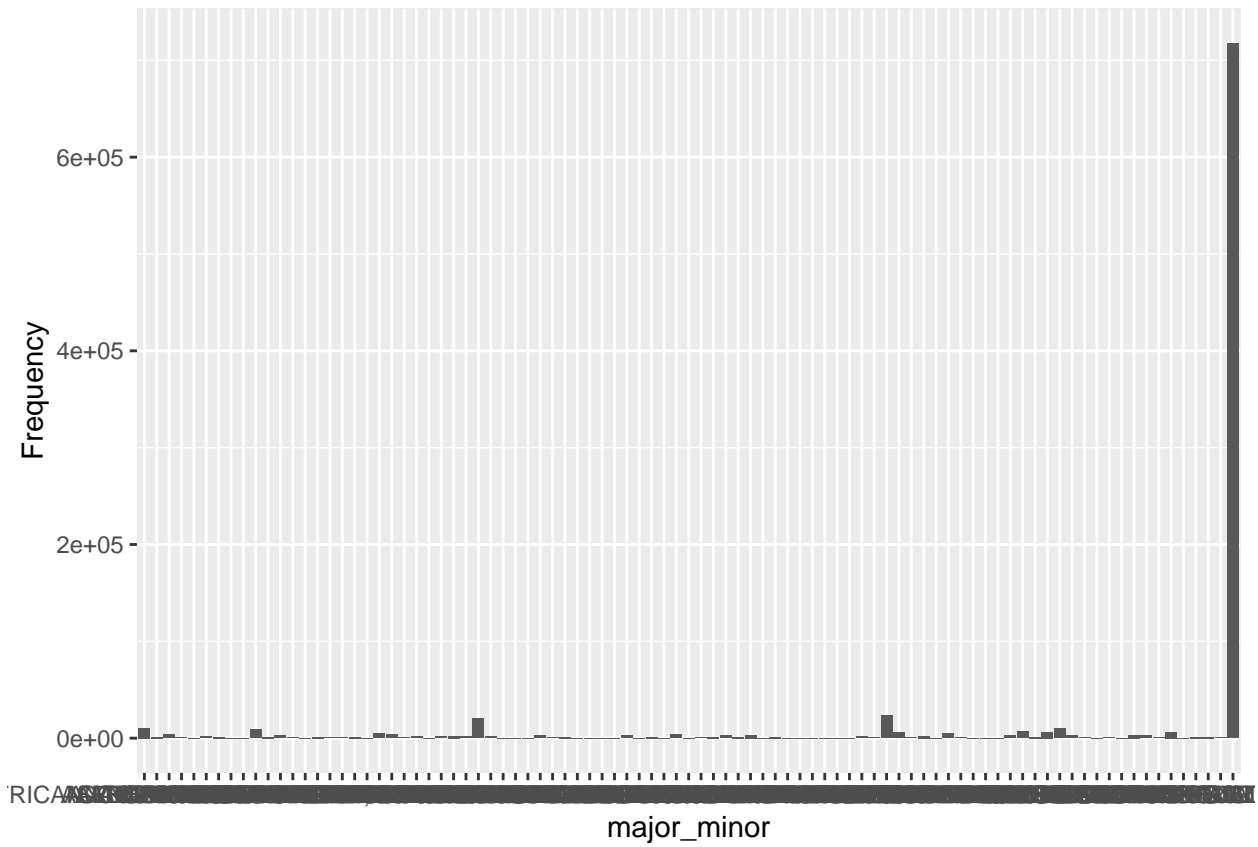
```
[1] -----  
[1] Variable: major_degree_granting_4, type: logical, requested: yes  
[1] Values (2 unique): NA, TRUE  
[1] Missing: 100%
```

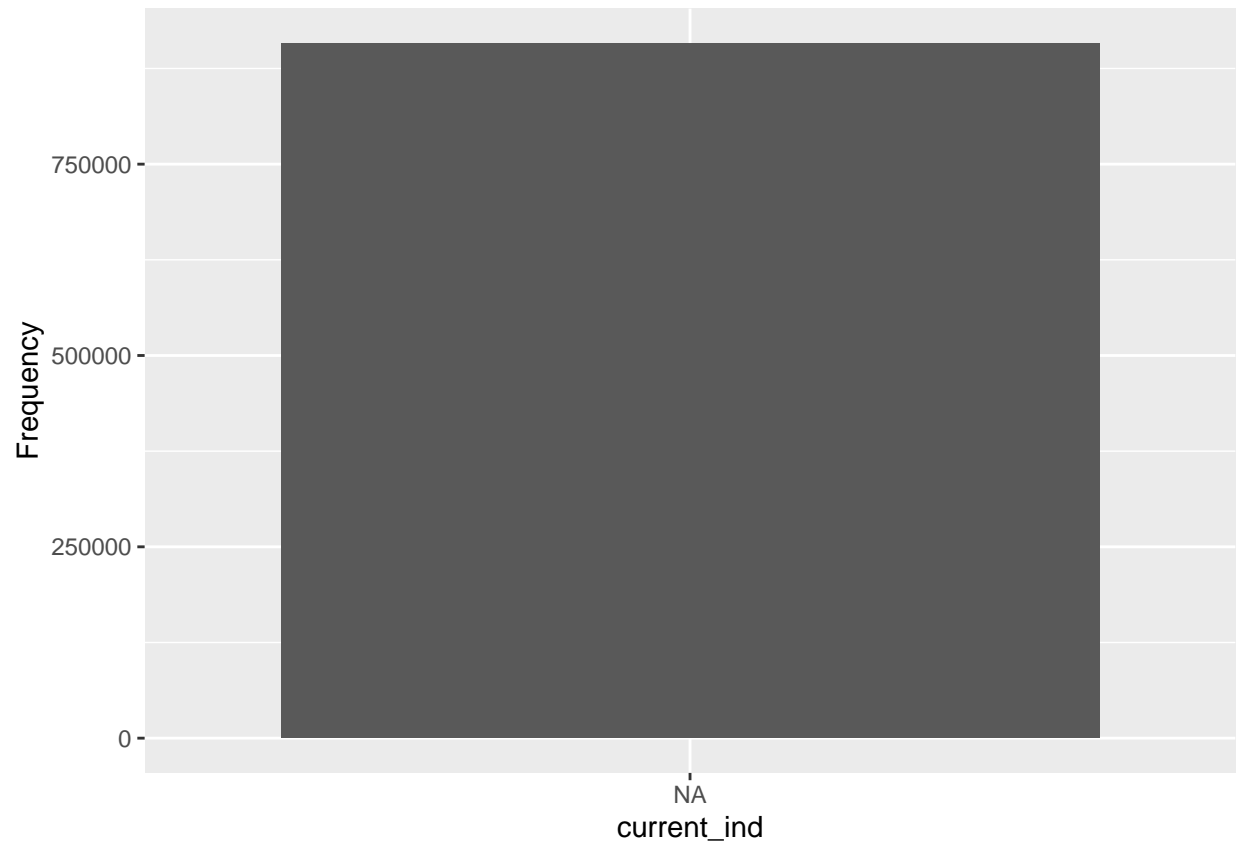
```
[1] -----  
[1] Variable: major_graduated_4, type: character, requested: yes  
[1] Values (1 unique): NA  
[1] Missing: 100%
```



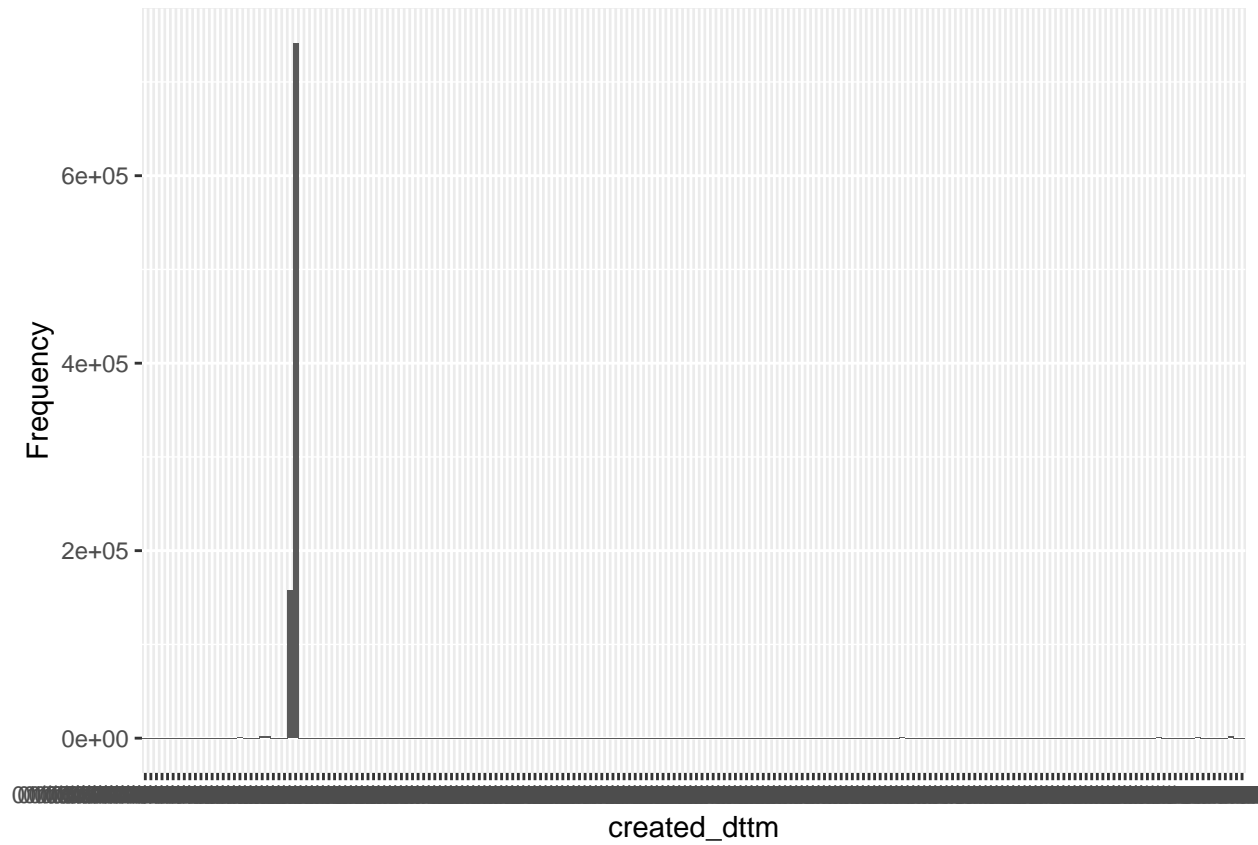
```
[1] -----
[1] Variable: major_minor, type: character, requested: yes
[1] Values (89 unique): NA, INNOVATION AND ENTREPRENEURSHIP, MANAGEMENT, AFRICAN-AMERICAN STUDIES, LING
[1] Missing: 79%
[1] Most missing: 2016 82.2%, Least missing: 2006 72.3%
```



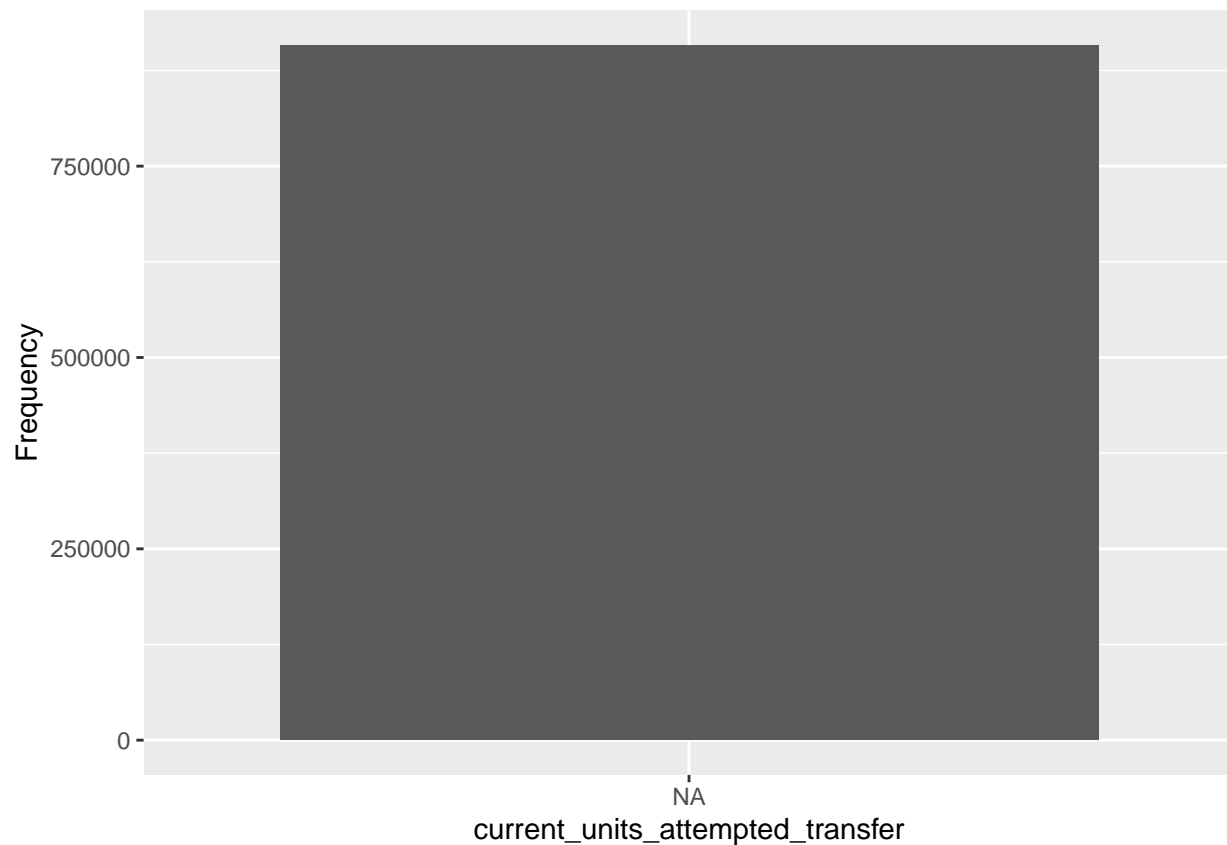
```
[1] -----
[1] Variable: current_ind, type: logical, requested: yes
[1] Values (1 unique): NA
[1] Missing: 100%
```



```
[1] -----
[1] Variable: created_dttm, type: character, requested: no
[1] Values (198 unique): 02/02/2023 08:55:38.000, 01/31/2023 09:19:06.000, 12/20/2022 09:18:06.000, 12/
[1] Missing: 0%
```

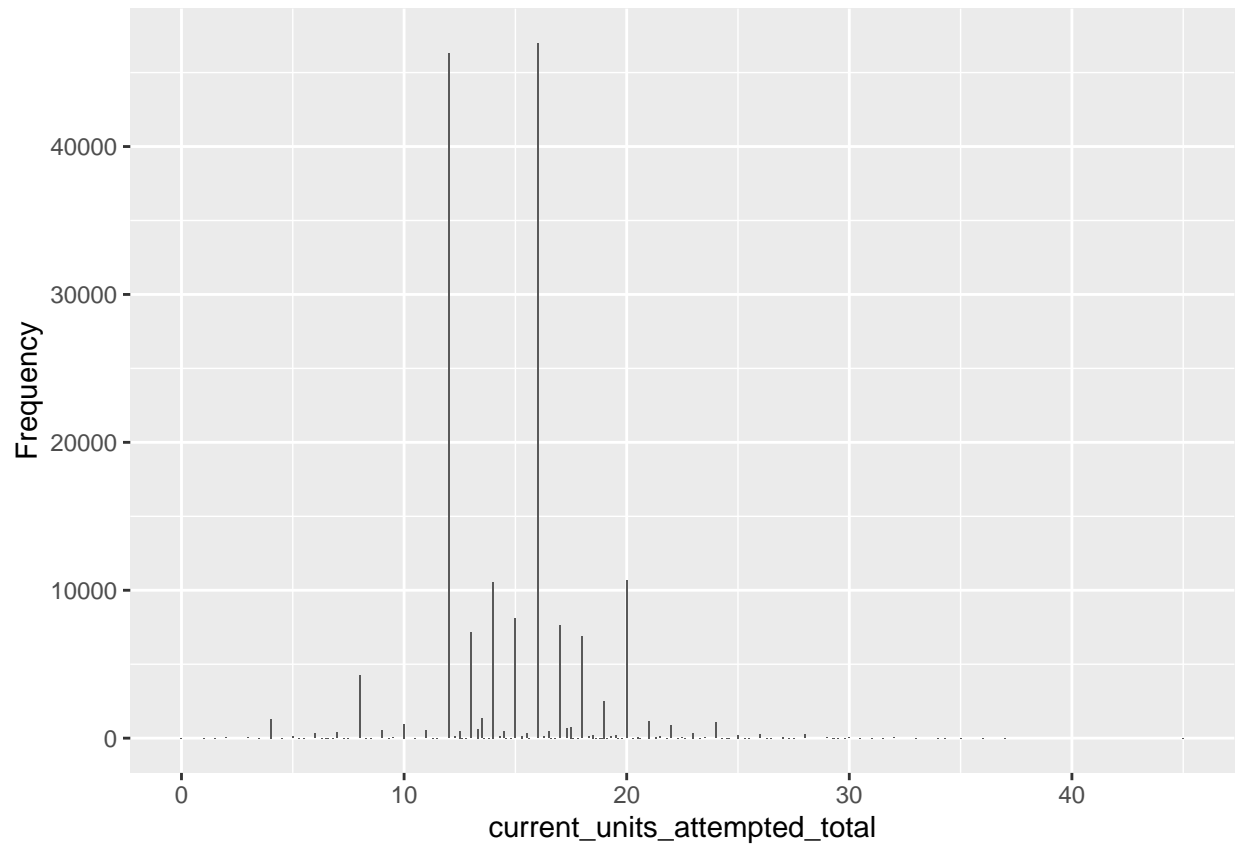


```
[1] -----  
[1] Variable: current_units_attempted_transfer, type: logical, requested: no  
[1] Values (1 unique): NA  
[1] Missing: 100%
```

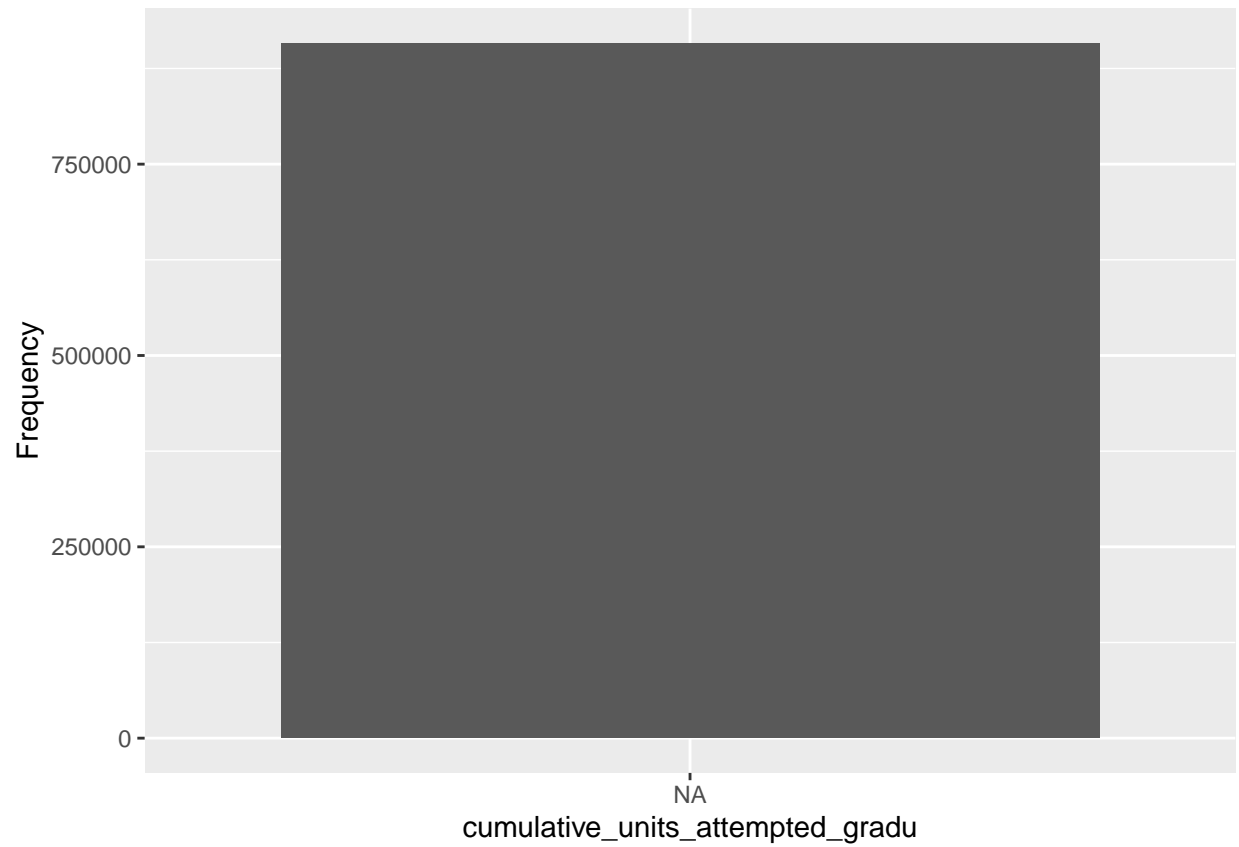


```
[1] -----
[1] Variable: current_units_attempted_total, type: numeric, requested: no
[1] Values (117 unique): NA, 12, 15, 14, 16, ...
[1] Missing: 81.7%
  Group.1      x
1   2006 0.9989400
2   2007 0.9995472
3   2008 0.9995467
4   2009 0.9992735
5   2010 0.9993125
6   2011 0.9980382
7   2012 0.9972956
8   2013 0.9945398
9   2014 0.9770521
10  2015 0.8585476
11  2016 0.6644875
12  2017 0.4200917
13  2018 0.1084650
```

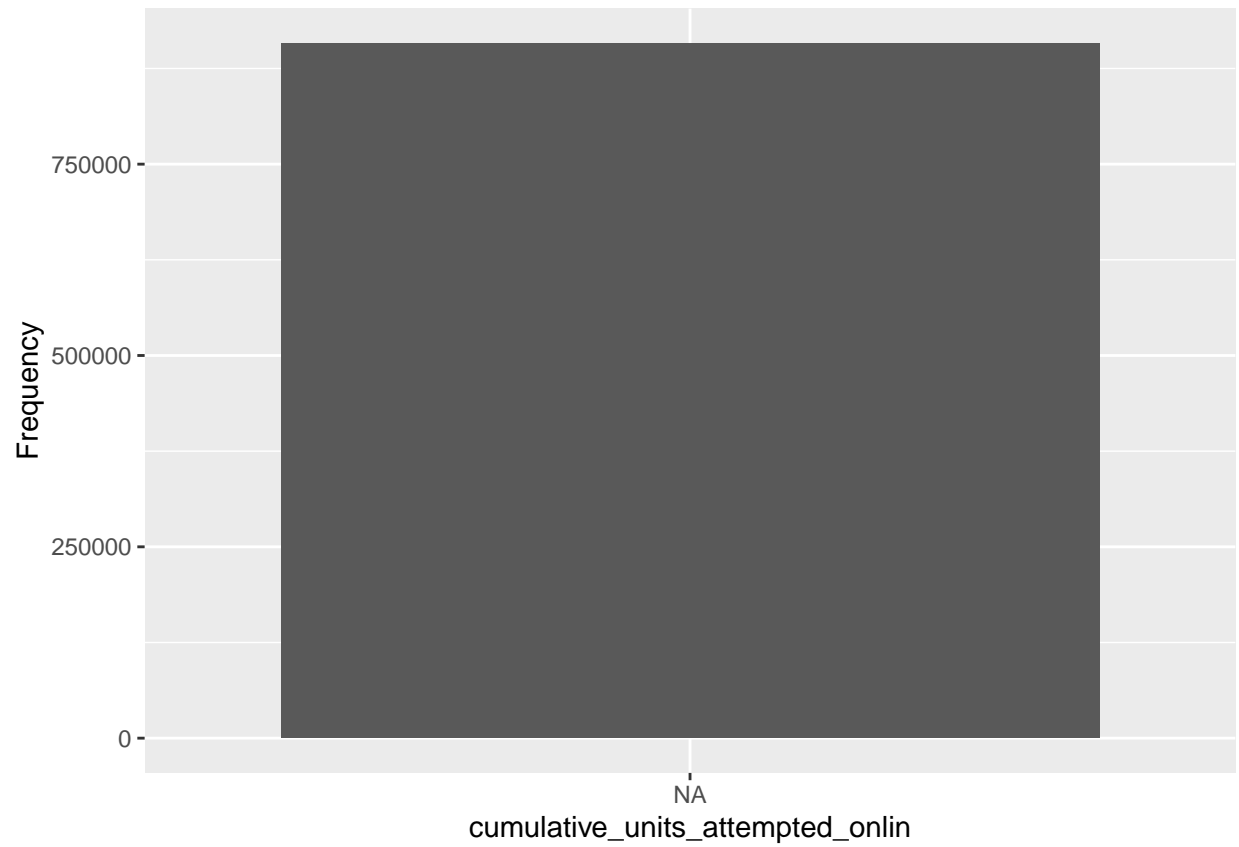
Warning: Removed 741555 rows containing non-finite values ('stat_count()').



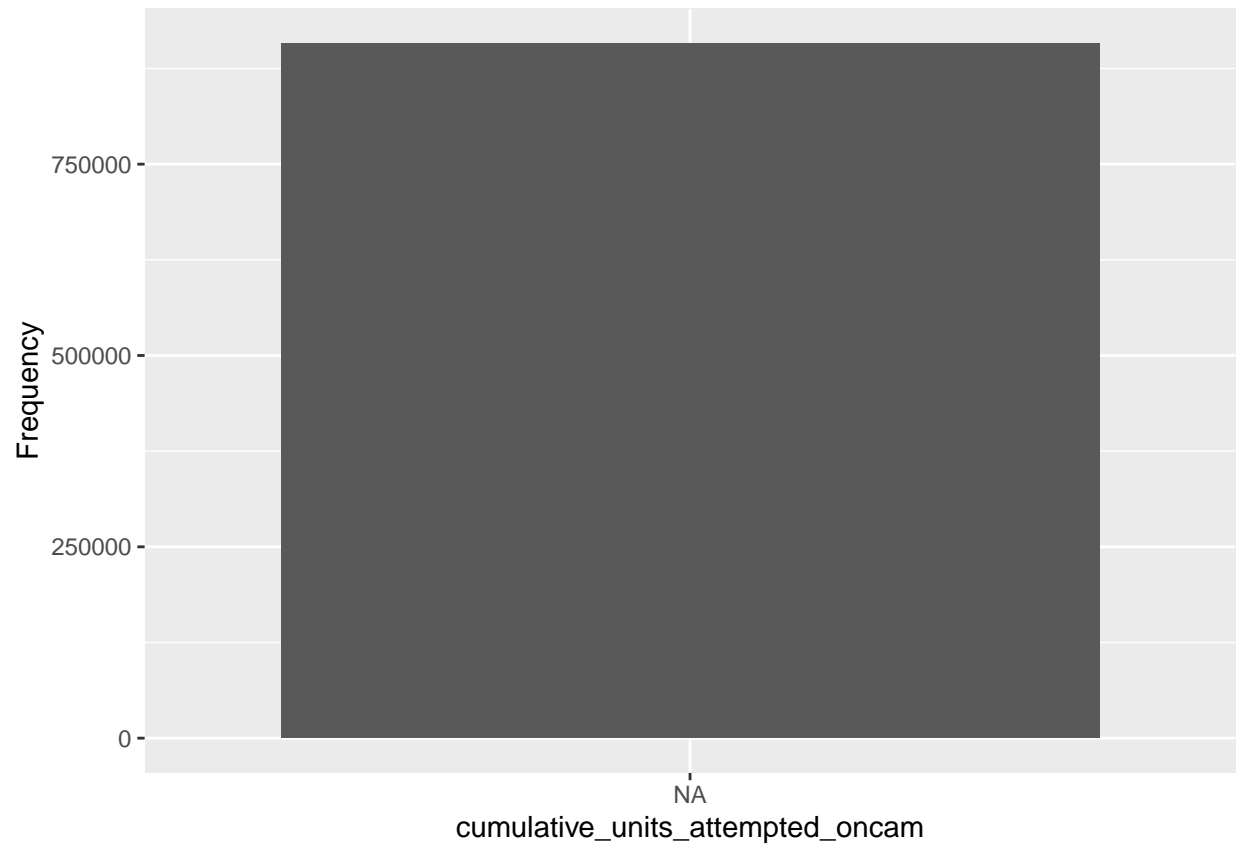
```
[1] -----  
[1] Variable: cumulative_units_attempted_gradu, type: logical, requested: no  
[1] Values (1 unique): NA  
[1] Missing: 100%
```



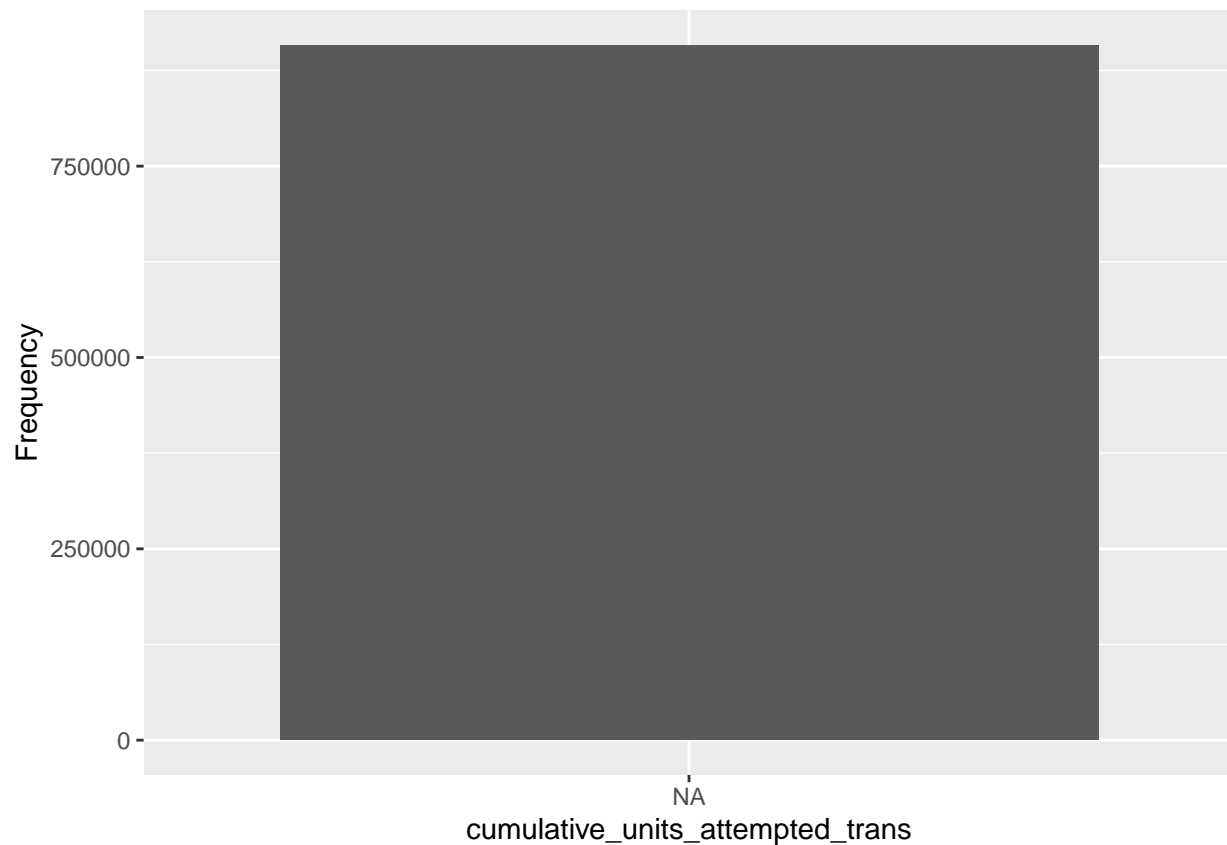
```
[1] -----  
[1] Variable: cumulative_units_attempted_onlin, type: logical, requested: no  
[1] Values (1 unique): NA  
[1] Missing: 100%
```

```
[1] -----  
[1] Variable: cumulative_units_attempted_oncam, type: logical, requested: no  
[1] Values (1 unique): NA  
[1] Missing: 100%
```

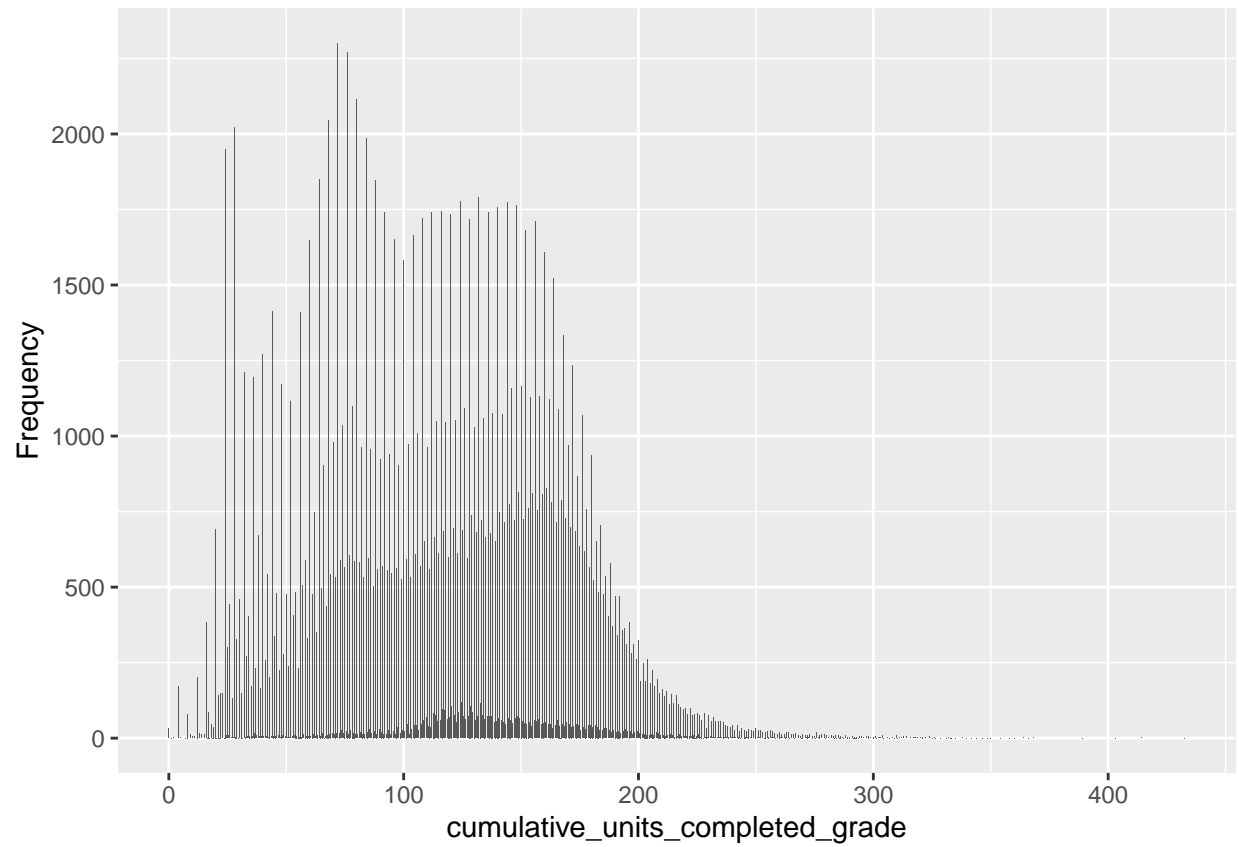


```
[1] -----  
[1] Variable: cumulative_units_attempted_trans, type: logical, requested: no  
[1] Values (1 unique): NA  
[1] Missing: 100%
```

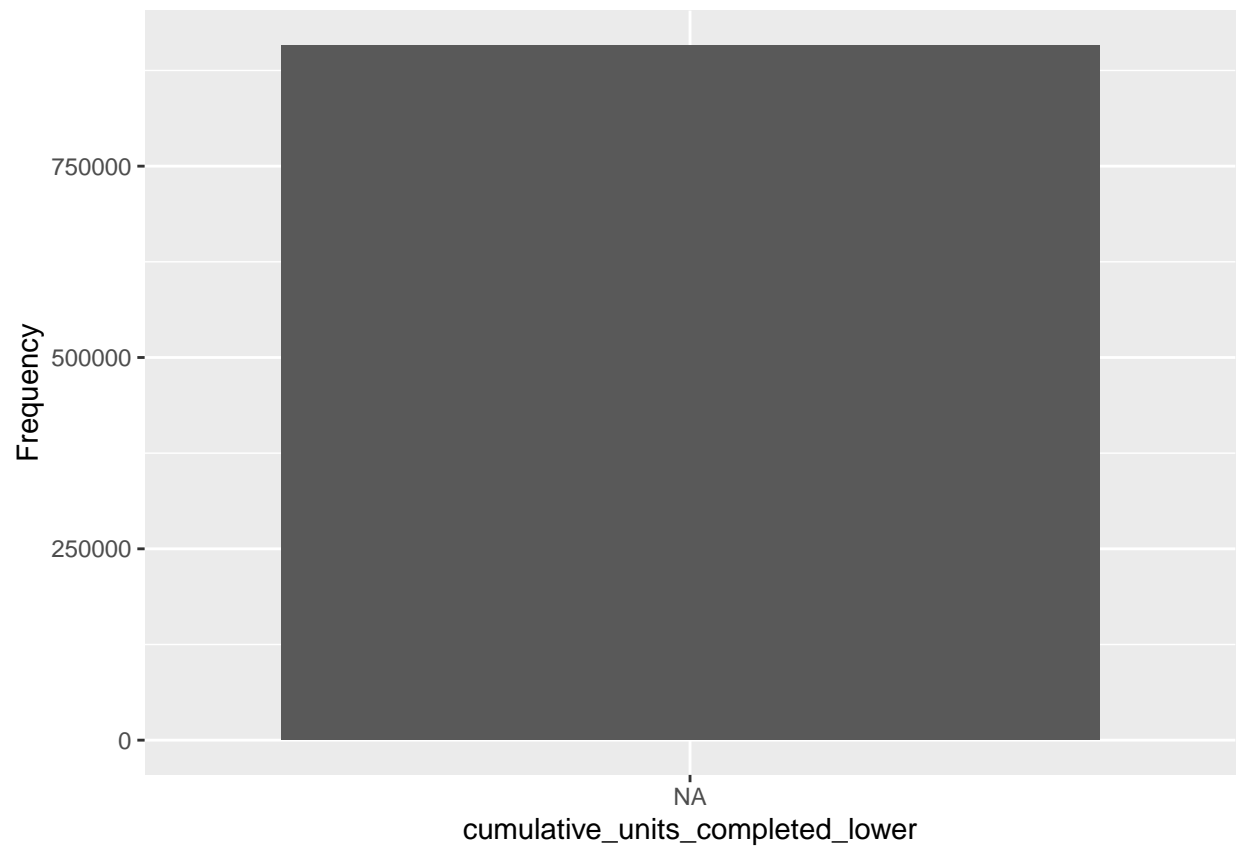


```
[1] -----
[1] Variable: cumulative_units_completed_grade, type: numeric, requested: no
[1] Values (625 unique): NA, 64, 79, 36, 151, ...
[1] Missing: 81.7%
  Group.1      x
1    2006 0.9989164
2    2007 0.9995310
3    2008 0.9995311
4    2009 0.9992562
5    2010 0.9993125
6    2011 0.9980093
7    2012 0.9972217
8    2013 0.9945128
9    2014 0.9770521
10   2015 0.8589157
11   2016 0.6668134
12   2017 0.4215572
13   2018 0.1076654
```

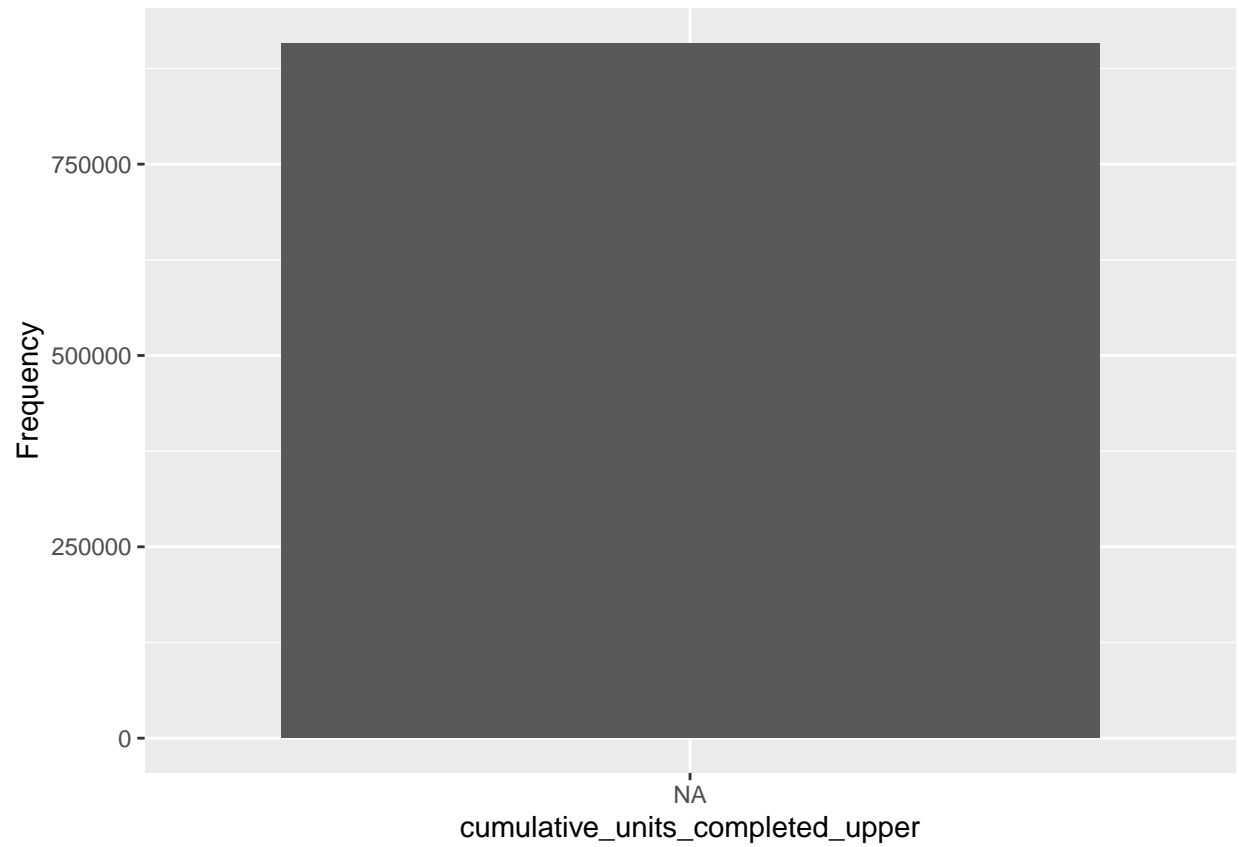
Warning: Removed 741843 rows containing non-finite values ('stat_count()').



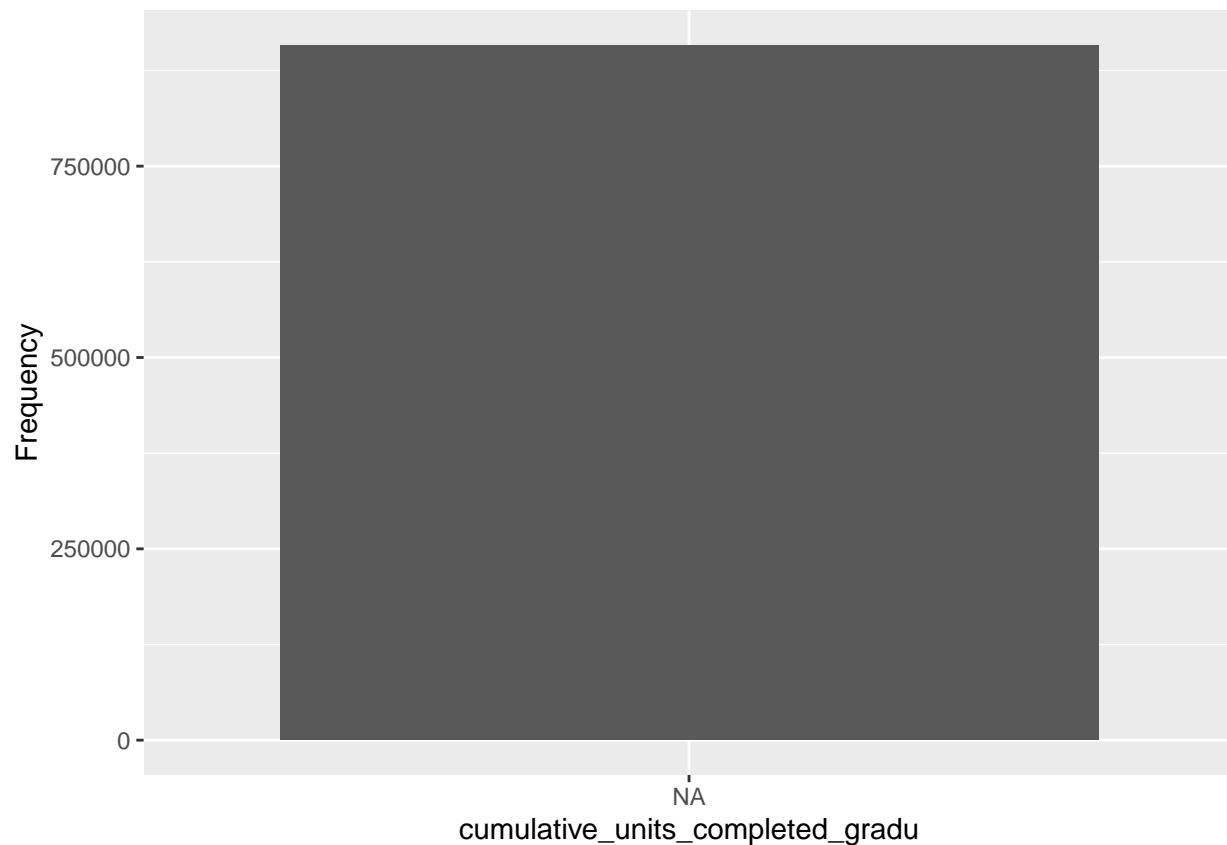
```
[1] -----  
[1] Variable: cumulative_units_completed_lower, type: logical, requested: no  
[1] Values (1 unique): NA  
[1] Missing: 100%
```



```
[1] -----  
[1] Variable: cumulative_units_completed_upper, type: logical, requested: no  
[1] Values (1 unique): NA  
[1] Missing: 100%
```

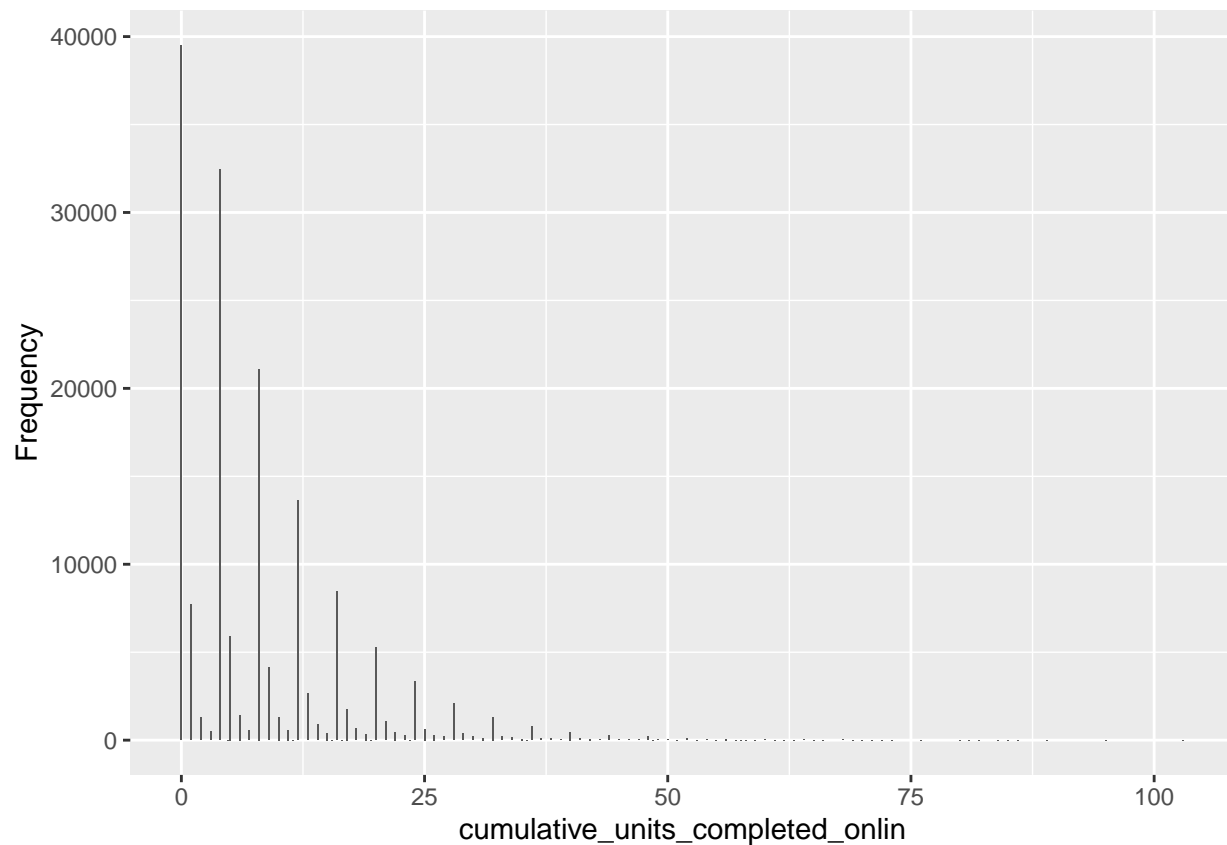


```
[1] -----  
[1] Variable: cumulative_units_completed_gradu, type: logical, requested: no  
[1] Values (1 unique): NA  
[1] Missing: 100%
```



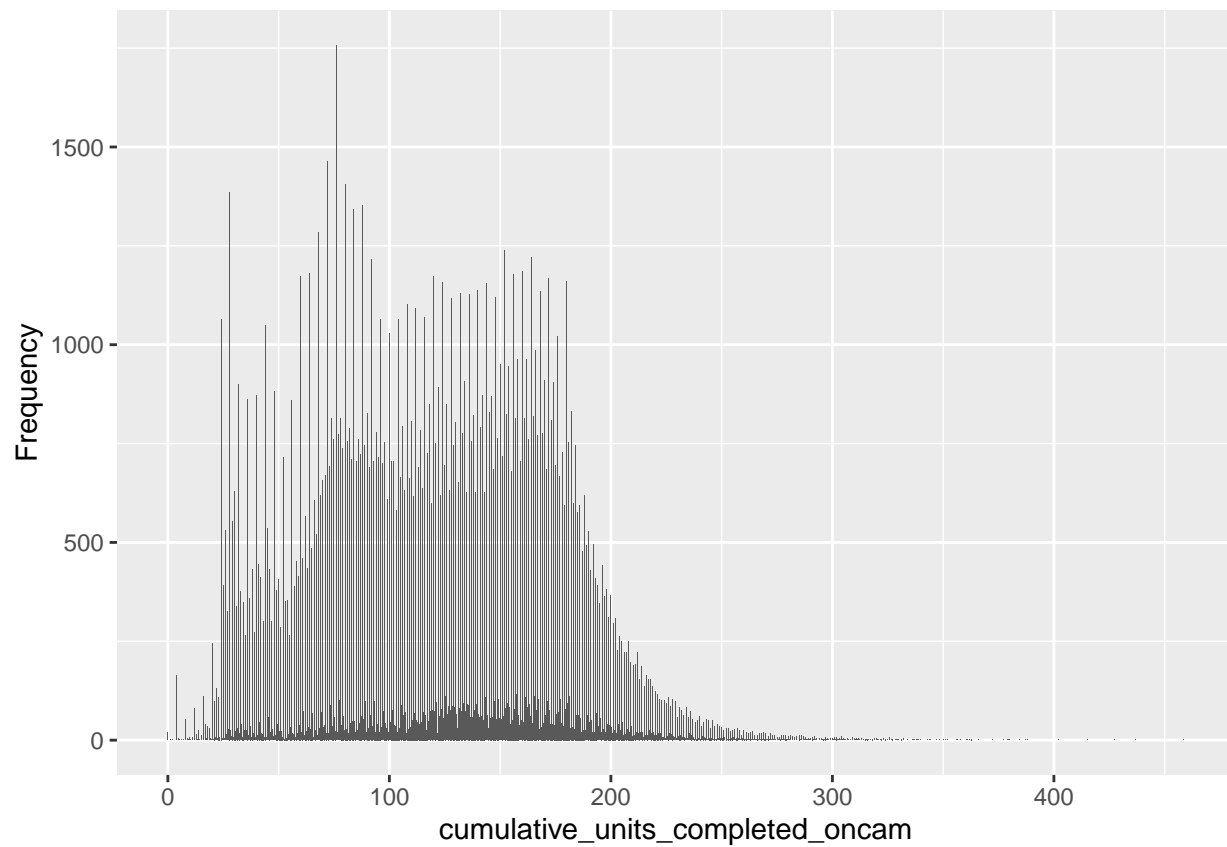
```
[1] -----
[1] Variable: cumulative_units_completed_onlin, type: numeric, requested: no
[1] Values (93 unique): NA, 8, 4, 0, 12, ...
[1] Missing: 81.9%
  Group.1      x
1    2006 0.9990813
2    2007 0.9995796
3    2008 0.9995623
4    2009 0.9992908
5    2010 0.9993285
6    2011 0.9981103
7    2012 0.9972808
8    2013 0.9946614
9    2014 0.9774431
10   2015 0.8598754
11   2016 0.6678296
12   2017 0.4245777
13   2018 0.1262698
```

Warning: Removed 743831 rows containing non-finite values ('stat_count()').

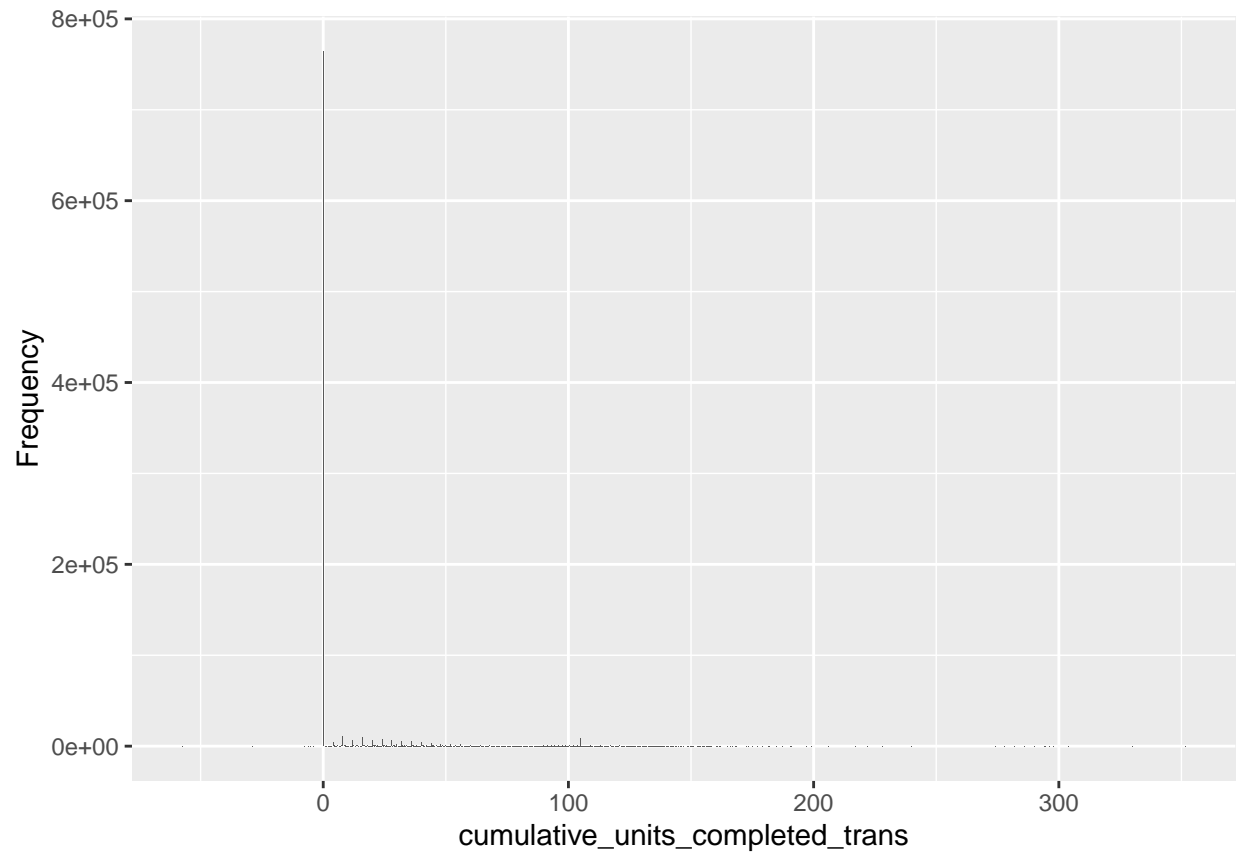


```
[1] -----
[1] Variable: cumulative_units_completed_oncam, type: numeric, requested: no
[1] Values (2070 unique): NA, 64, 79, 36, 151, ...
[1] Missing: 81.7%
  Group.1      x
1    2006 0.9989164
2    2007 0.9995310
3    2008 0.9995311
4    2009 0.9992562
5    2010 0.9993125
6    2011 0.9980093
7    2012 0.9972217
8    2013 0.9945128
9    2014 0.9770521
10   2015 0.8589157
11   2016 0.6668134
12   2017 0.4215572
13   2018 0.1076654
```

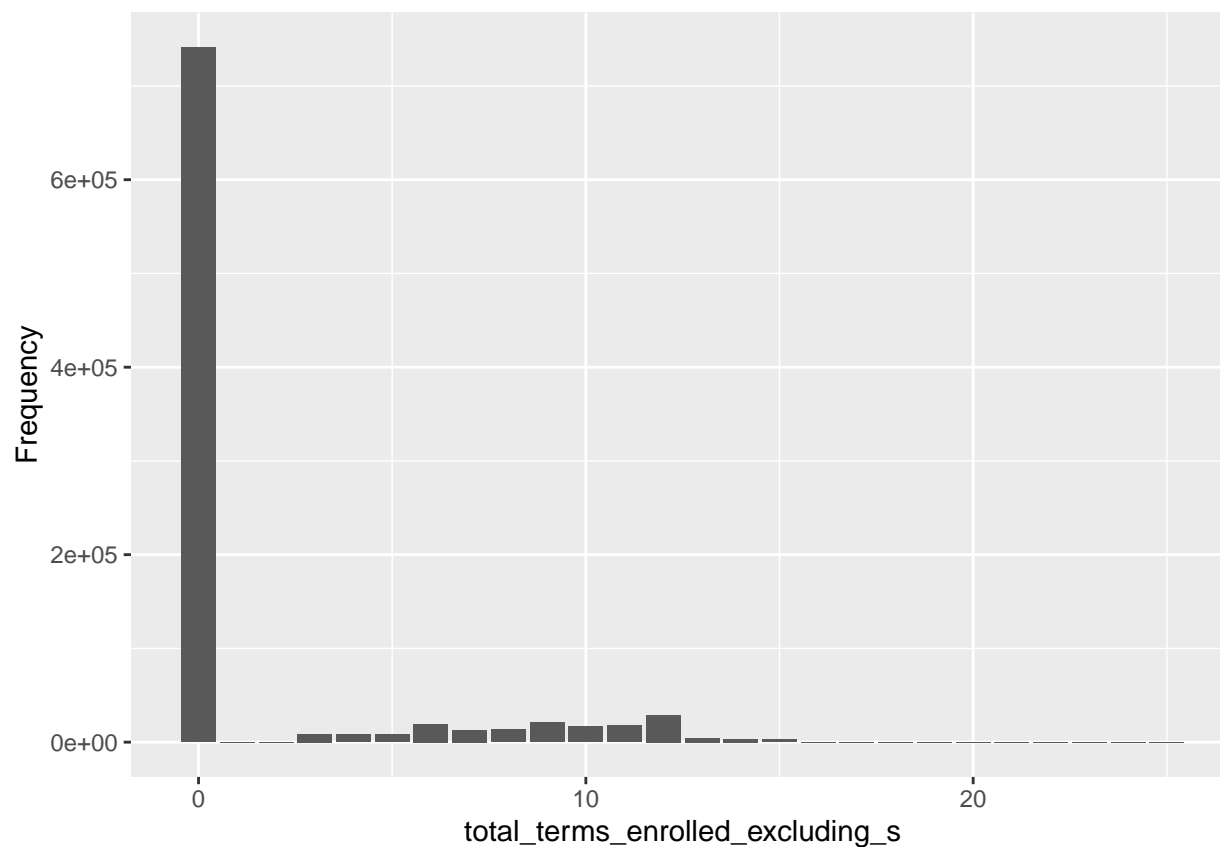
Warning: Removed 741843 rows containing non-finite values ('stat_count()').



```
[1] -----  
[1] Variable: cumulative_units_completed_trans, type: numeric, requested: no  
[1] Values (538 unique): 0, 32, 20, 16, 105, ...  
[1] Missing: 0%
```

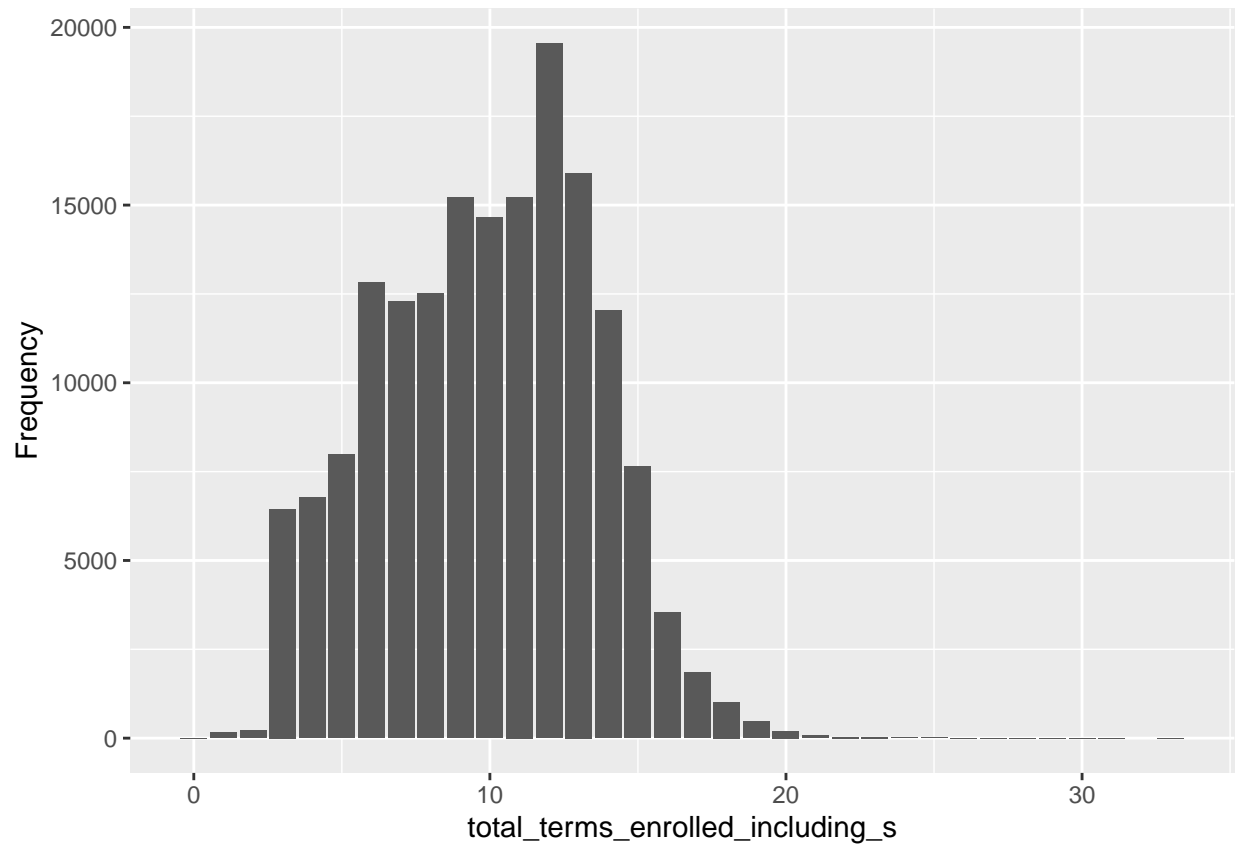


```
[1] -----  
[1] Variable: total_terms_enrolled_excluding_s, type: numeric, requested: no  
[1] Values (26 unique): 0, 6, 7, 4, 12, ...  
[1] Missing: 0%
```

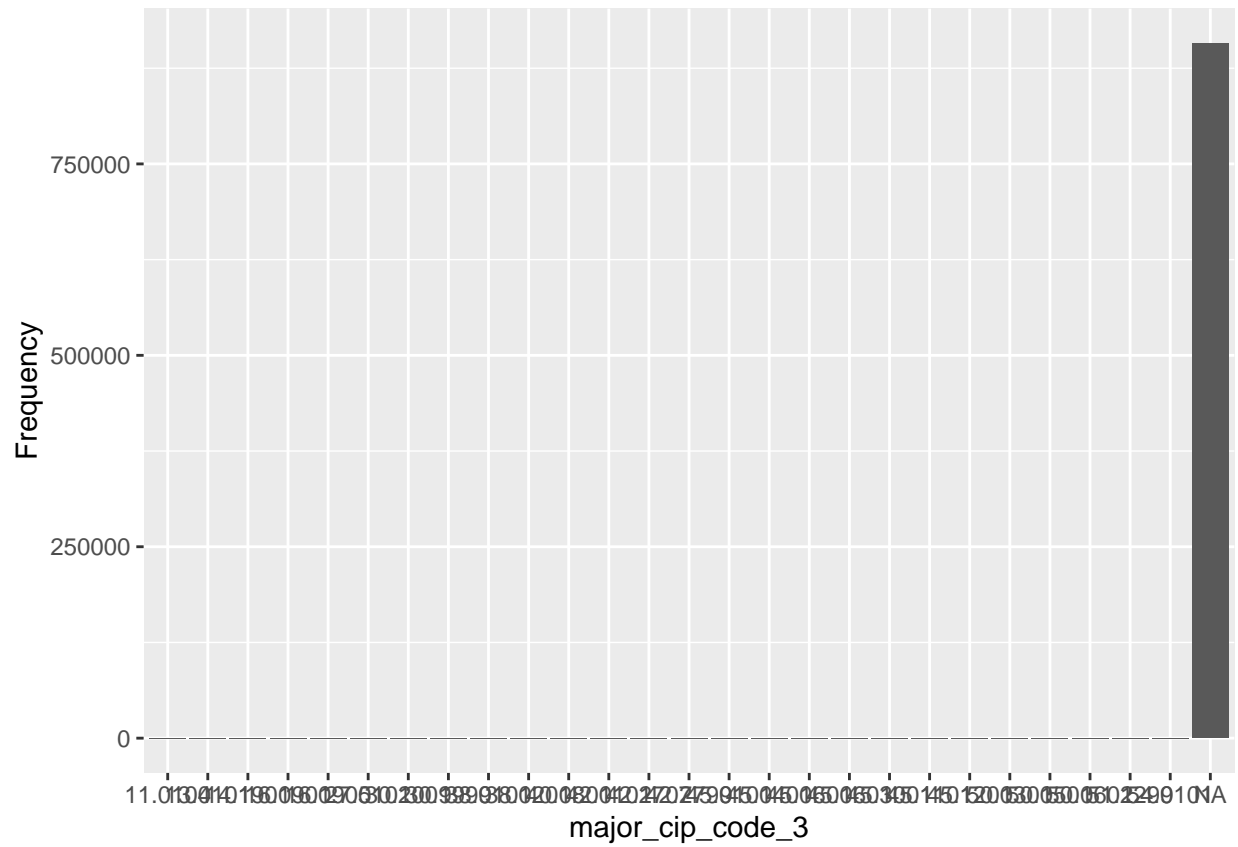


```
[1] -----
[1] Variable: total_terms_enrolled_including_s, type: numeric, requested: no
[1] Values (34 unique): NA, 8, 9, 5, 14, ...
[1] Missing: 81.6%
  Group.1      x
1    2006 0.9989164
2    2007 0.9995310
3    2008 0.9995311
4    2009 0.9992562
5    2010 0.9993125
6    2011 0.9980093
7    2012 0.9972217
8    2013 0.9944993
9    2014 0.9769173
10   2015 0.8583504
11   2016 0.6640472
12   2017 0.4191185
13   2018 0.1058037
```

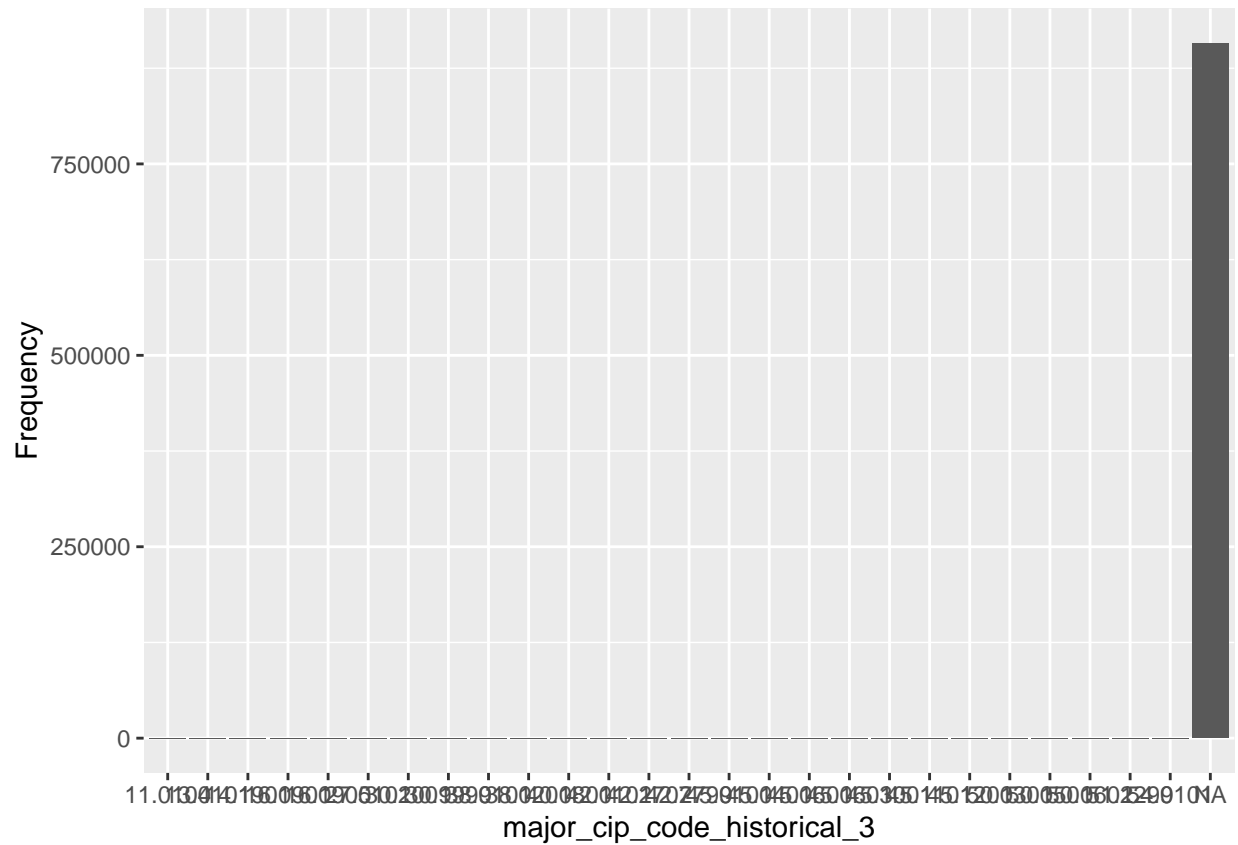
Warning: Removed 741177 rows containing non-finite values ('stat_count()').



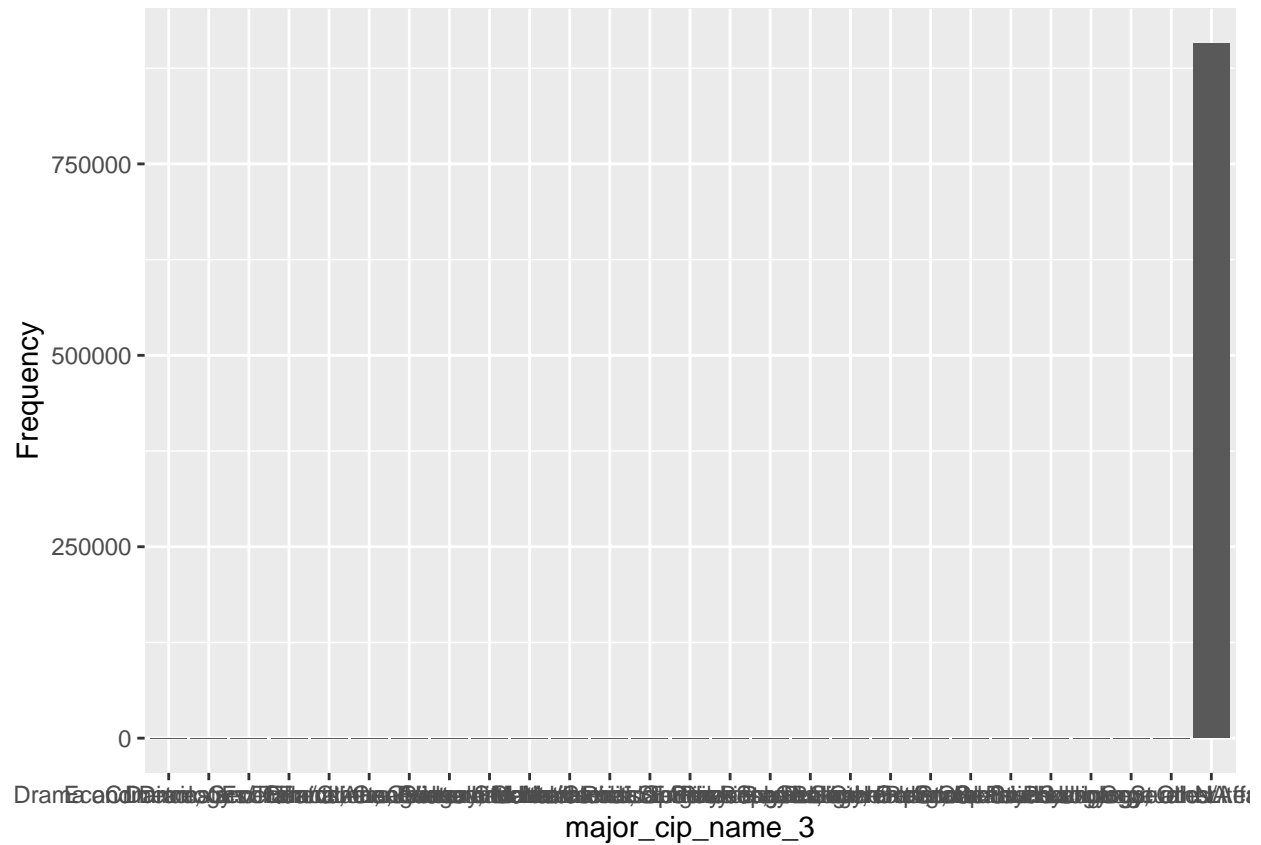
```
[1] -----  
[1] Variable: major_cip_code_3, type: character, requested: no  
[1] Values (27 unique): NA, 14.1901, 30.9999, 45.1101, 45.0101, ...  
[1] Missing: 100%
```



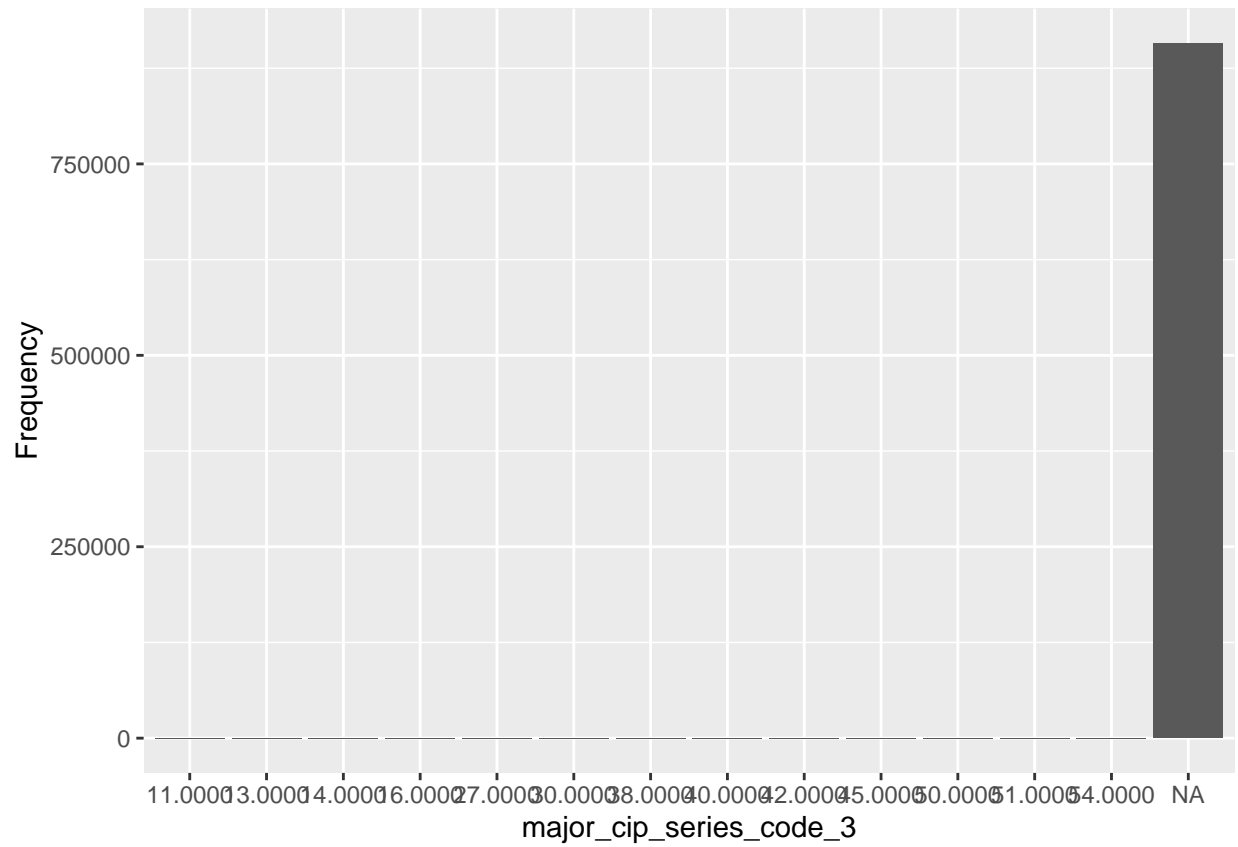
```
[1] -----
[1] Variable: major_cip_code_historical_3, type: character, requested: no
[1] Values (27 unique): NA, 14.1901, 30.9999, 45.1101, 45.0101, ...
[1] Missing: 100%
```



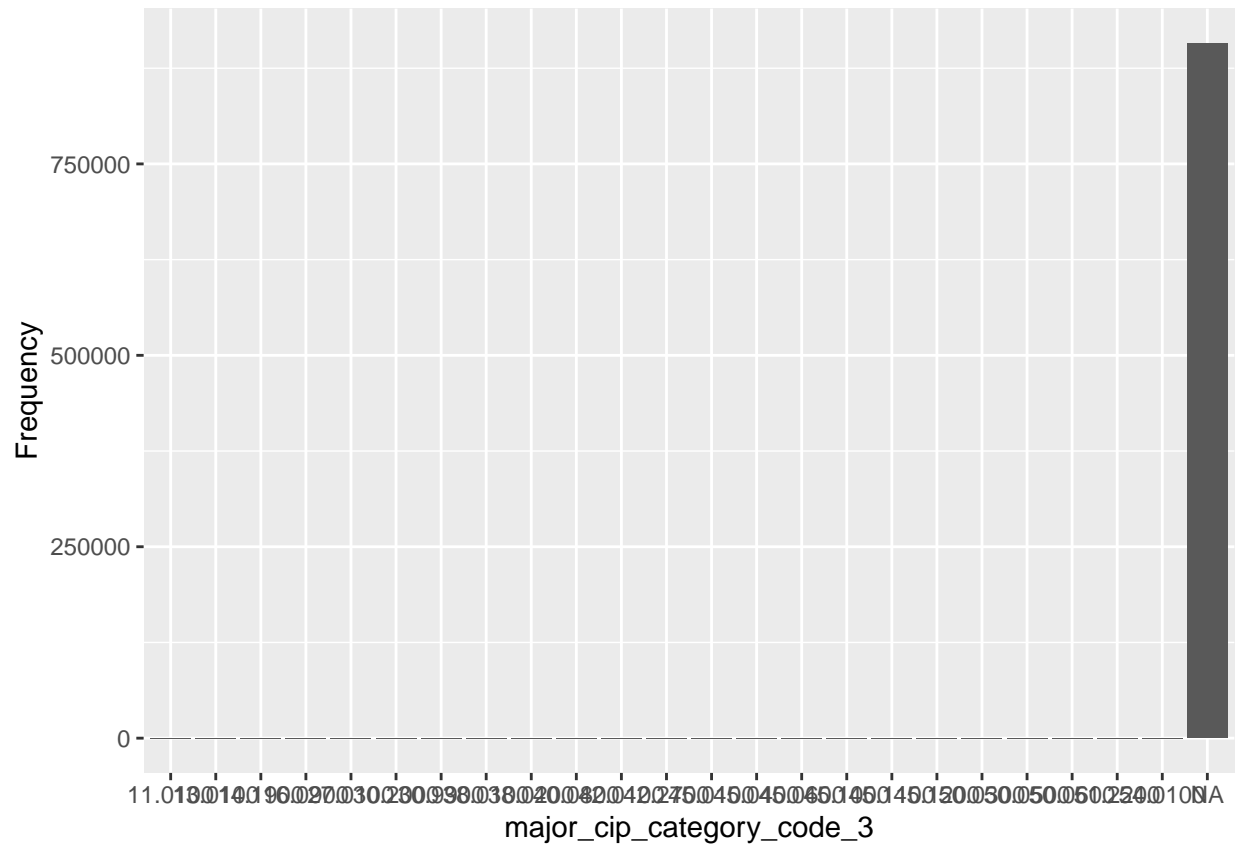
```
[1] -----
[1] Variable: major_cip_name_3, type: character, requested: no
[1] Values (27 unique): NA, Mechanical Engineering, Multi-/Interdisciplinary Studies, Other, Sociology,
[1] Missing: 100%
```



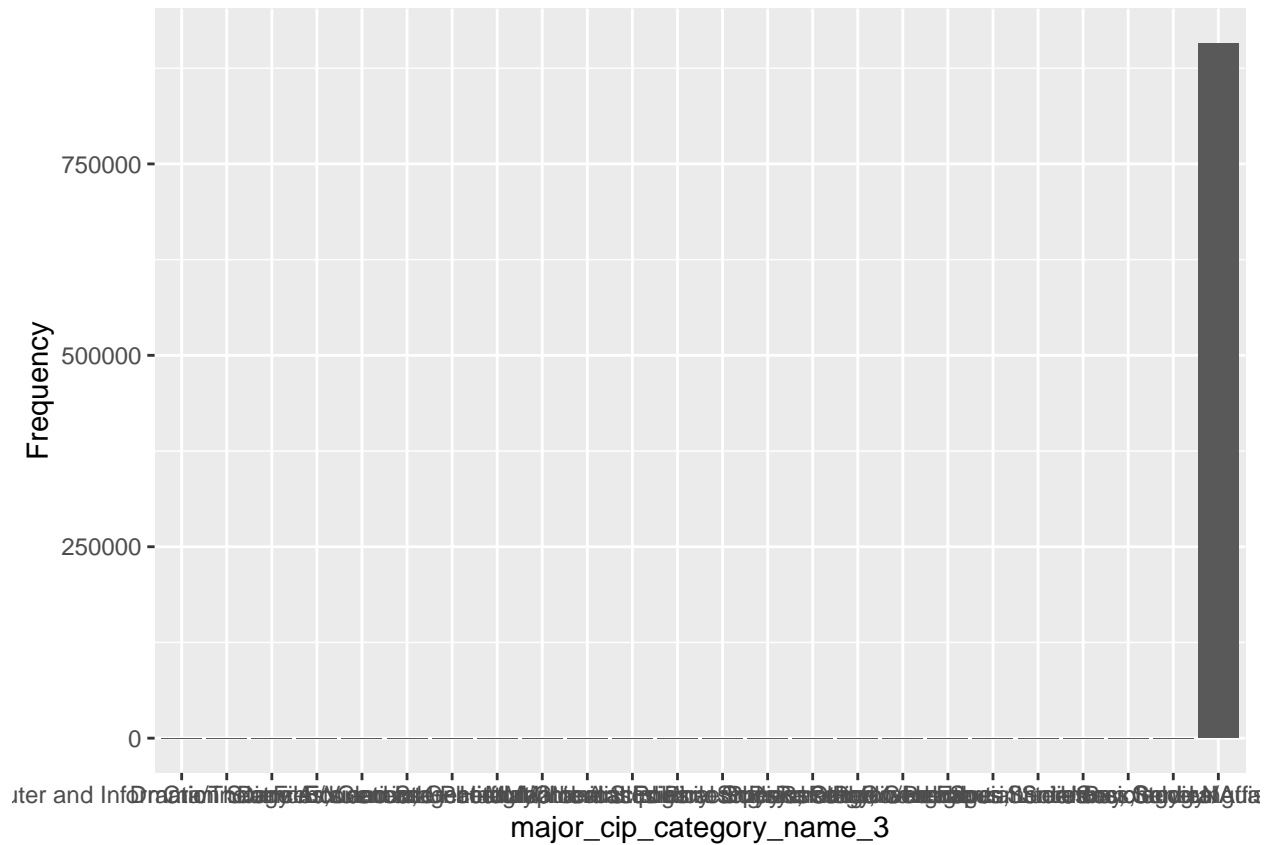
```
[1] -----
[1] Variable: major_cip_series_code_3, type: character, requested: no
[1] Values (14 unique): NA, 14.0000, 30.0000, 45.0000, 42.0000, ...
[1] Missing: 100%
```

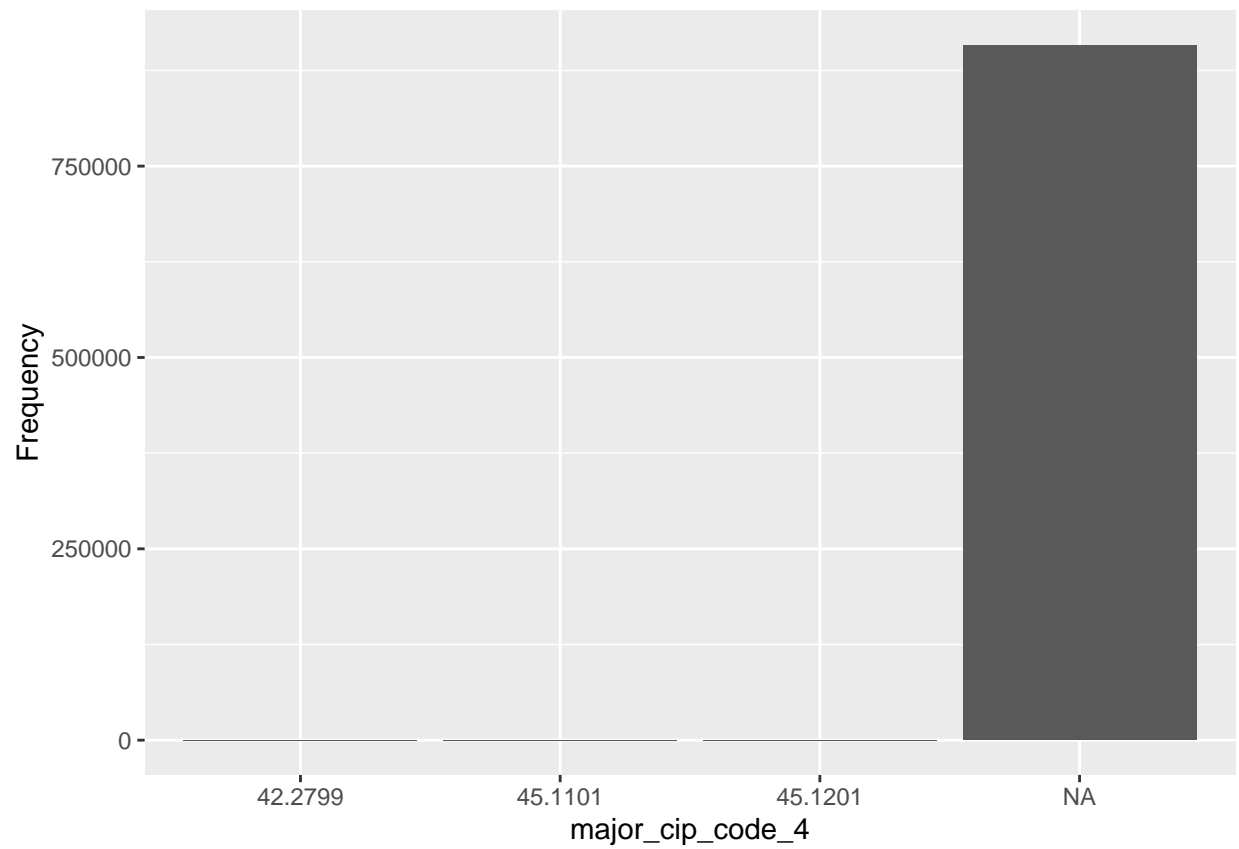


```
[1] -----
[1] Variable: major_cip_series_name_3, type: character, requested: no
[1] Values (14 unique): NA, Engineering, Multi/Interdisciplinary Studies, Social Sciences, Psychology,
[1] Missing: 100%
```

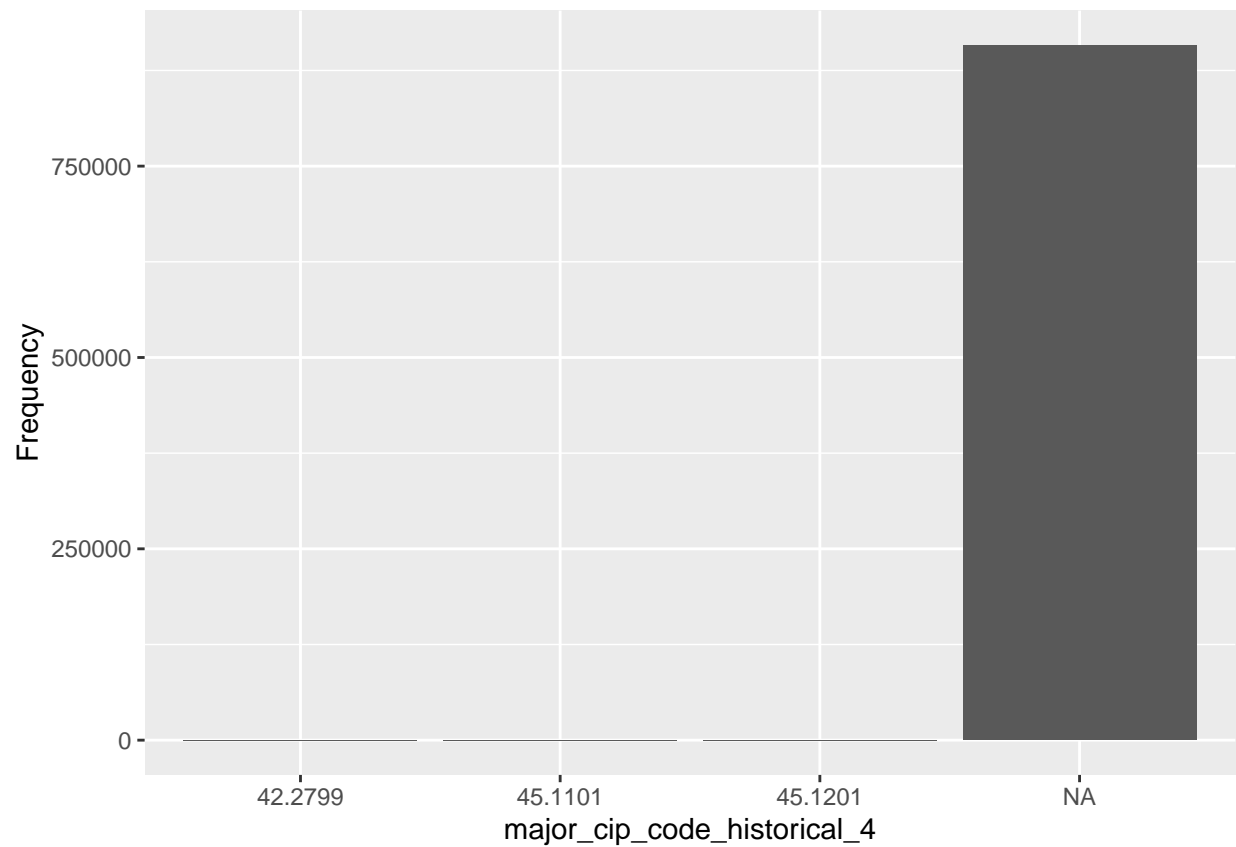



```
[1] -----
[1] Variable: major_cip_category_name_3, type: character, requested: no
[1] Values (24 unique): NA, Mechanical Engineering, Multi/Interdisciplinary Studies, Other, Sociology, S
[1] Missing: 100%
```

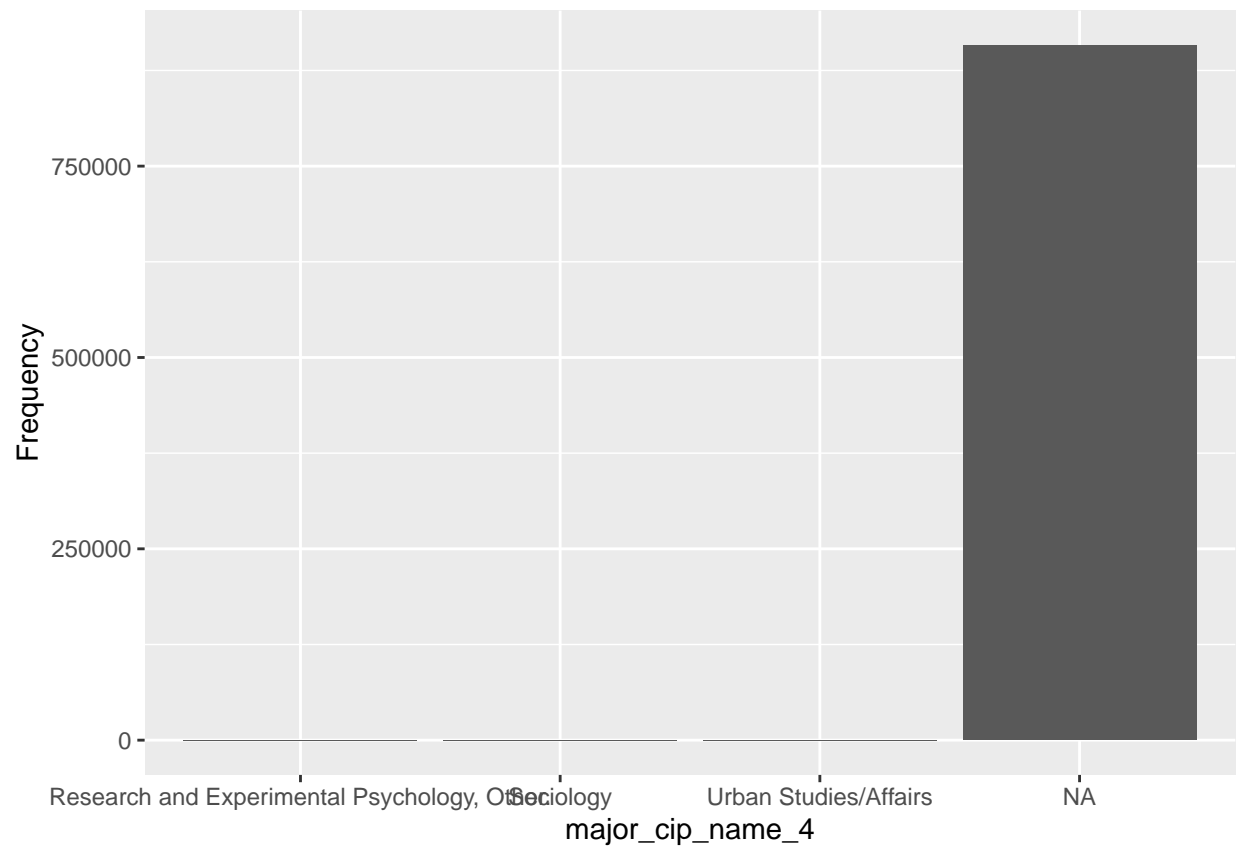




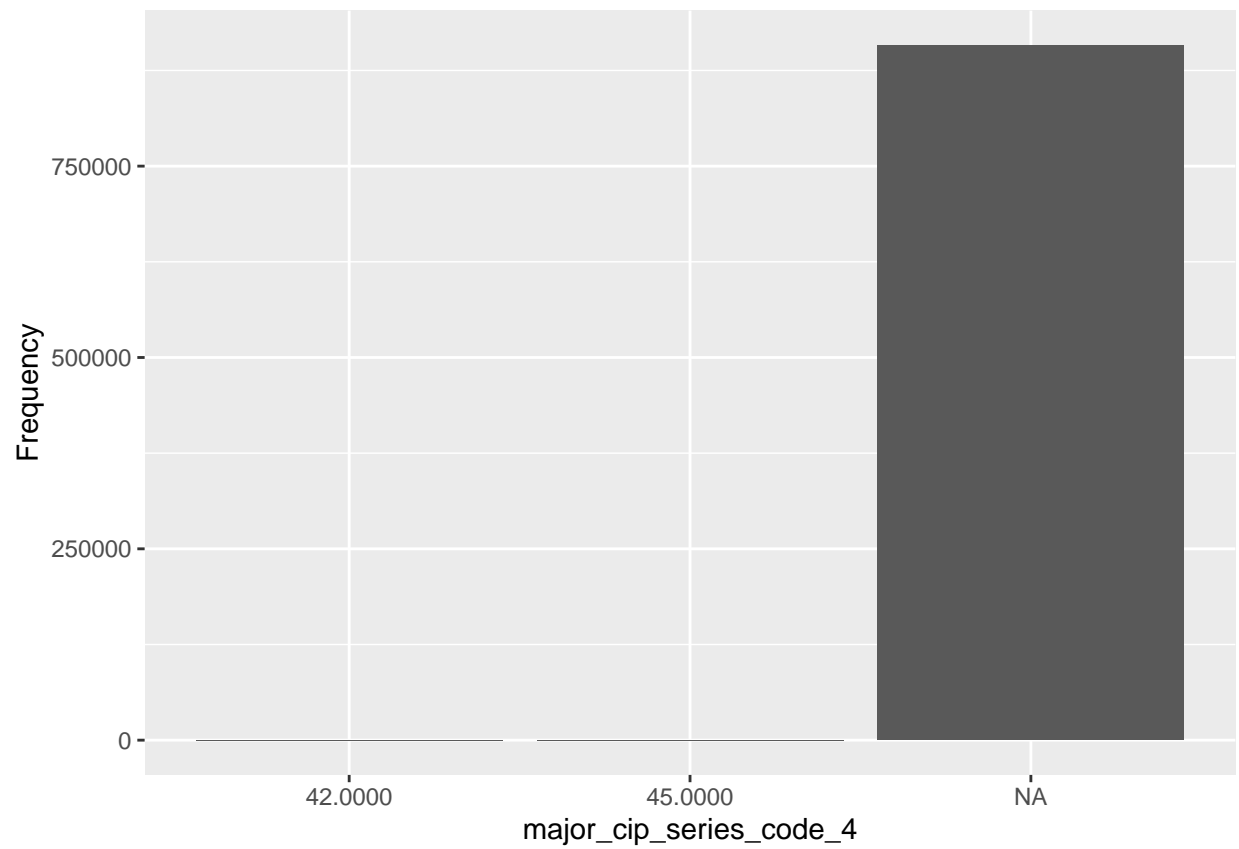
```
[1] -----  
[1] Variable: major_cip_code_historical_4, type: character, requested: no  
[1] Values (4 unique): NA, 45.1201, 42.2799, 45.1101  
[1] Missing: 100%
```



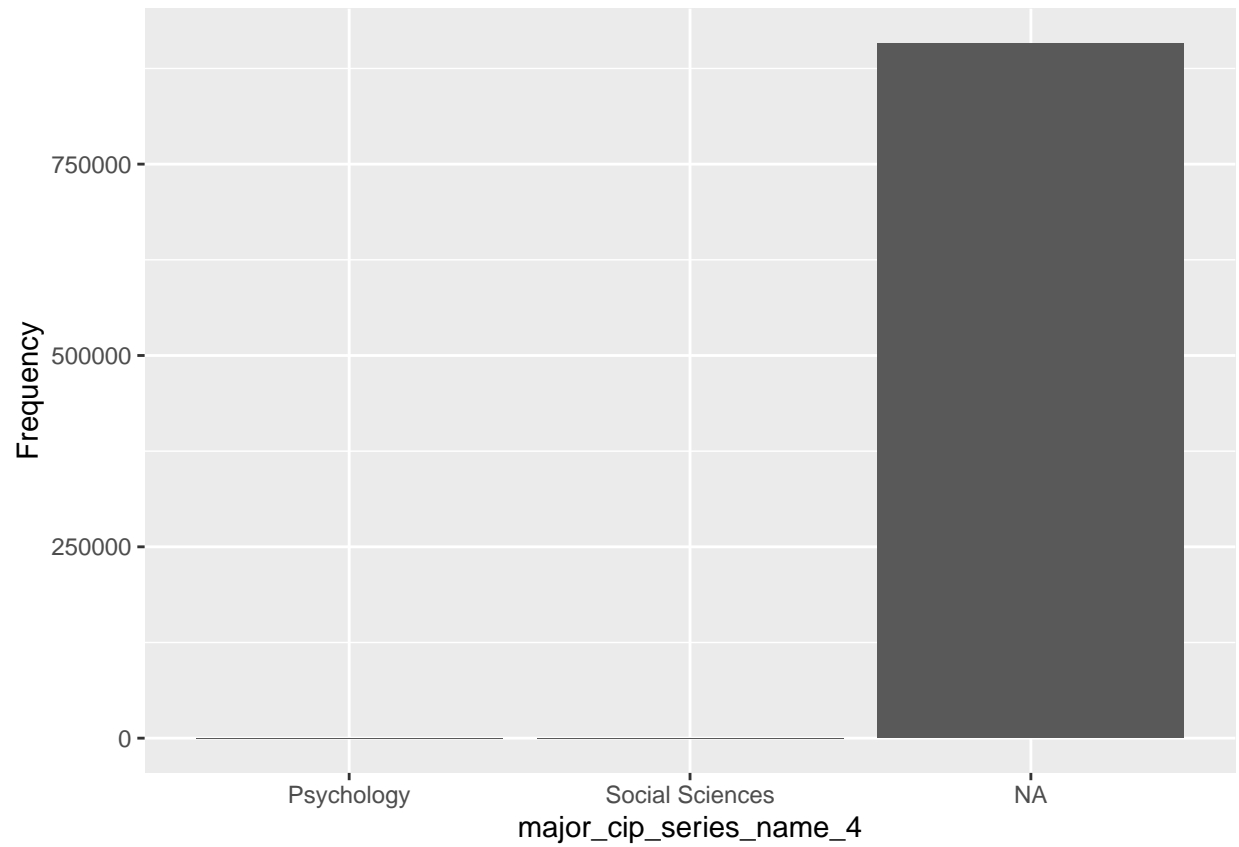
```
[1] -----
[1] Variable: major_cip_name_4, type: character, requested: no
[1] Values (4 unique): NA, Urban Studies/Affairs, Research and Experimental Psychology, Other., Sociology
[1] Missing: 100%
```



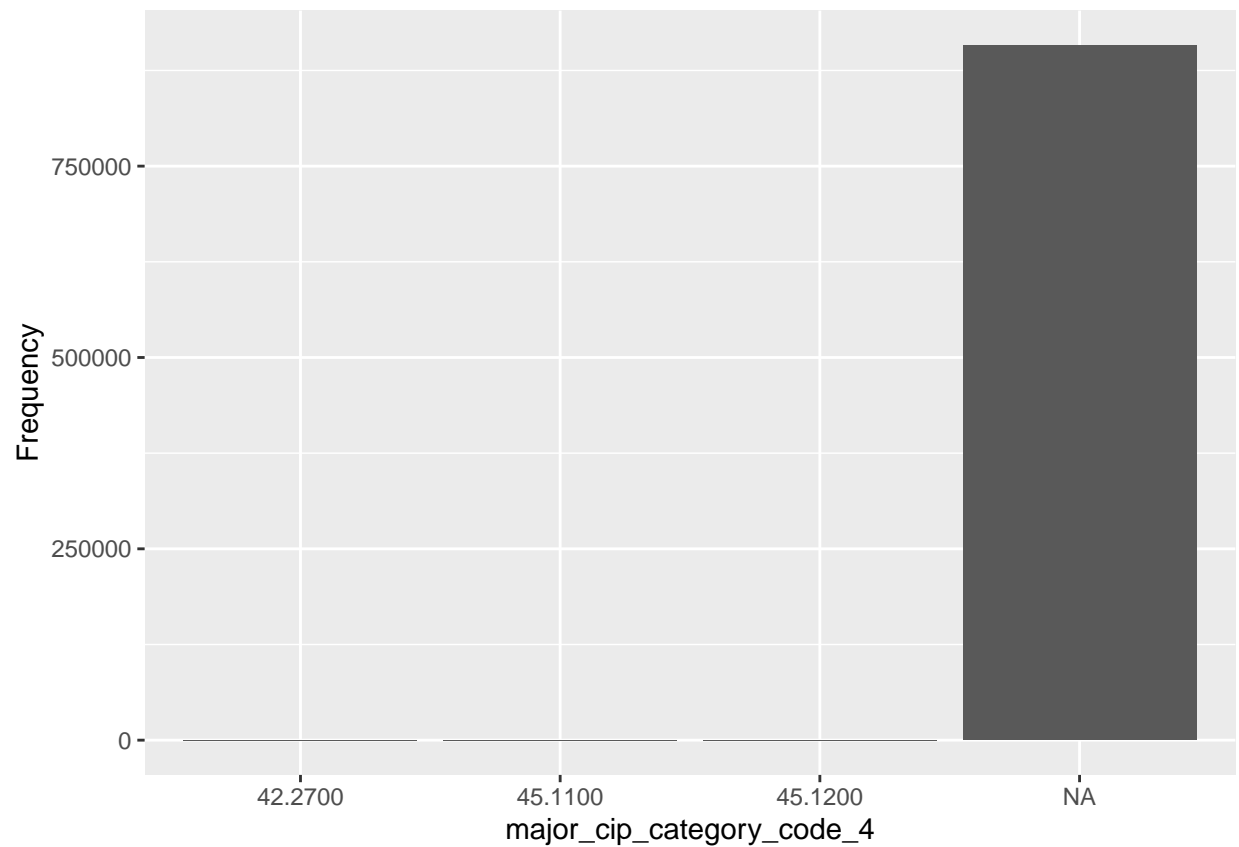
```
[1] -----
[1] Variable: major_cip_series_code_4, type: character, requested: no
[1] Values (3 unique): NA, 45.0000, 42.0000
[1] Missing: 100%
```



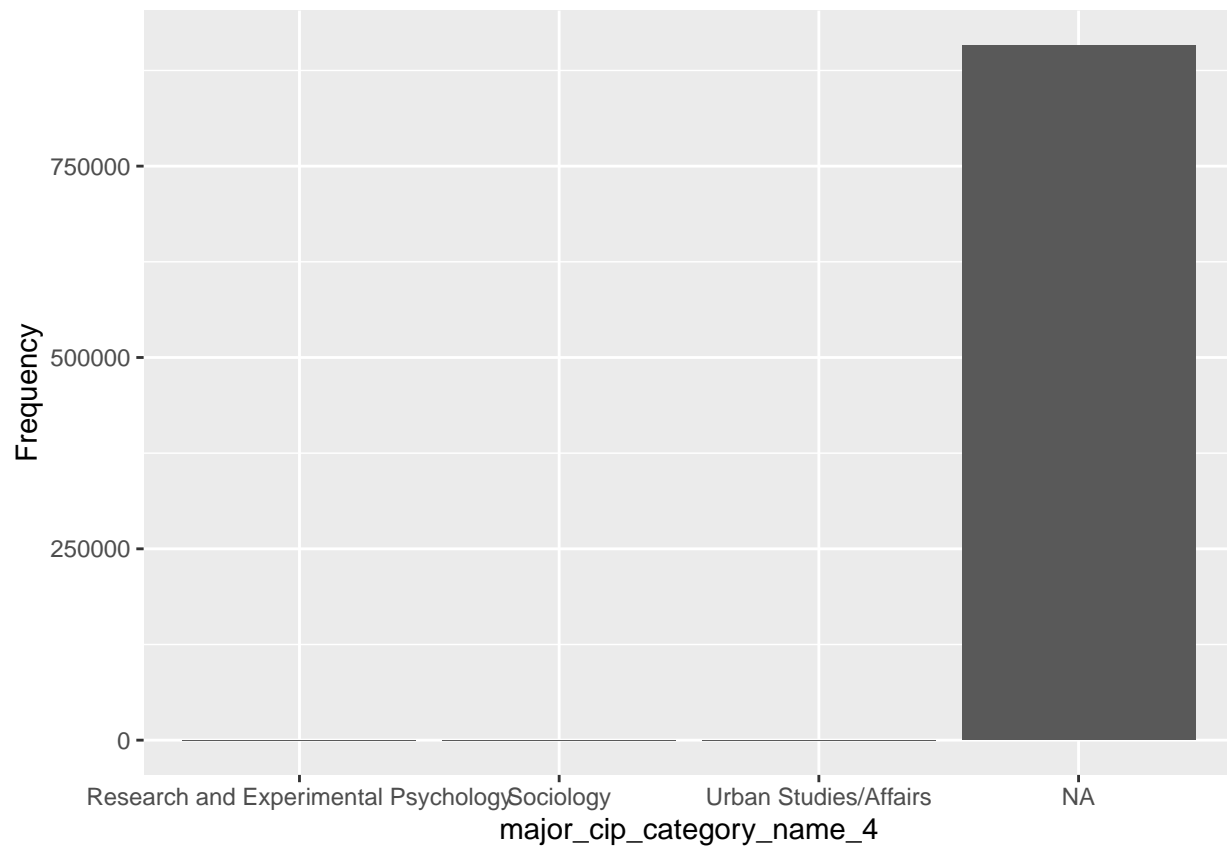
```
[1] -----  
[1] Variable: major_cip_series_name_4, type: character, requested: no  
[1] Values (3 unique): NA, Social Sciences, Psychology  
[1] Missing: 100%
```



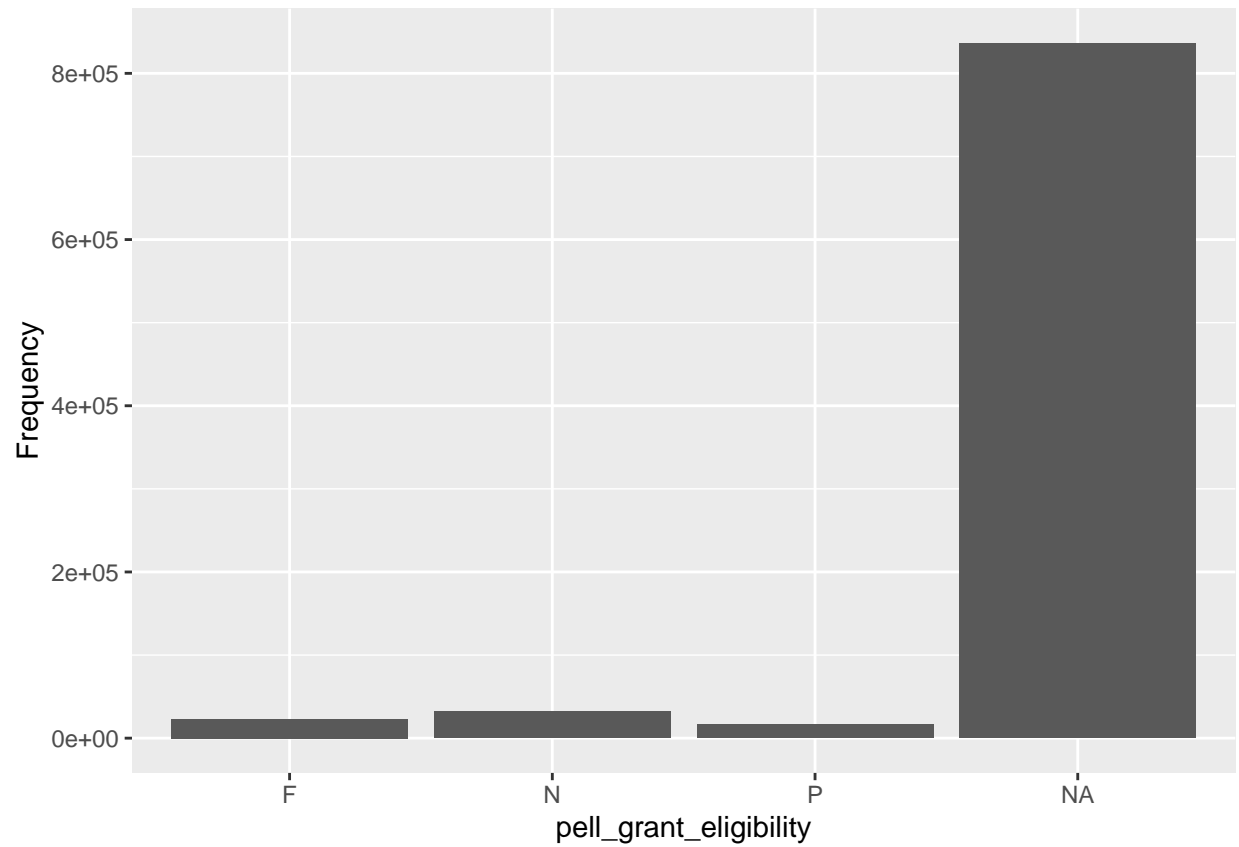
```
[1] -----  
[1] Variable: major_cip_category_code_4, type: character, requested: no  
[1] Values (4 unique): NA, 45.1200, 42.2700, 45.1100  
[1] Missing: 100%
```

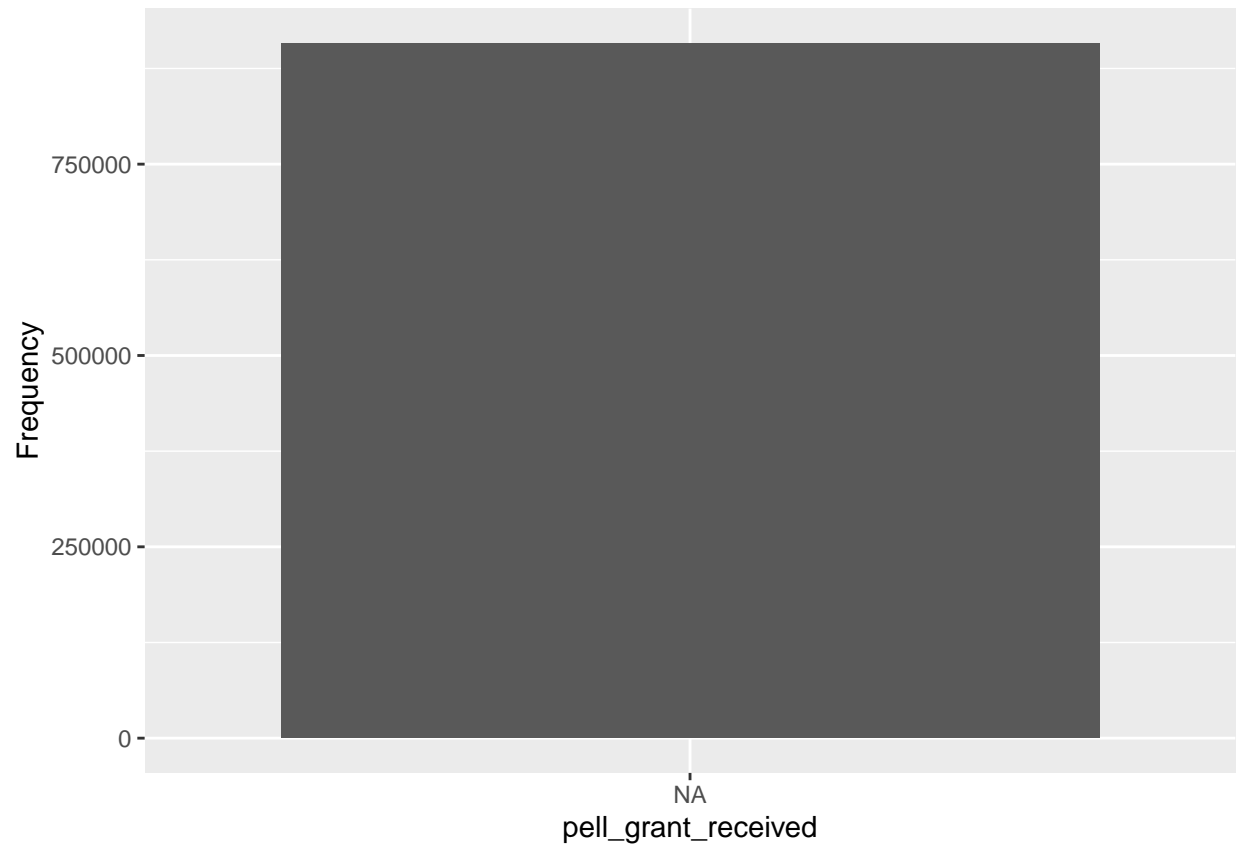
```
[1] -----
[1] Variable: major_cip_category_name_4, type: character, requested: no
[1] Values (4 unique): NA, Urban Studies/Affairs, Research and Experimental Psychology, Sociology
[1] Missing: 100%
```



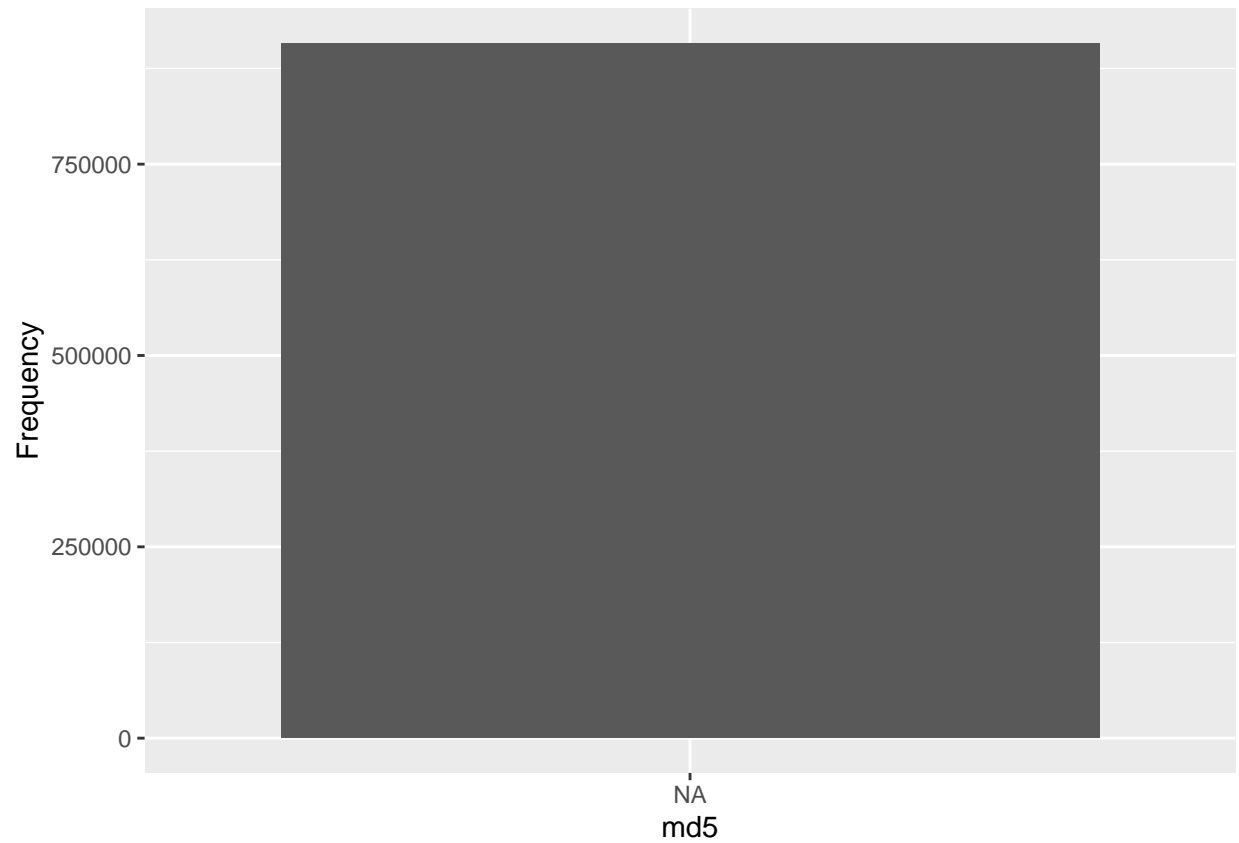
```
[1] -----
[1] Variable: pell_grant_eligibility, type: character, requested: no
[1] Values (4 unique): NA, N, P, F
[1] Missing: 92.1%
[1] Most missing: 2017 100%, Least missing: 2014 70.5%
```



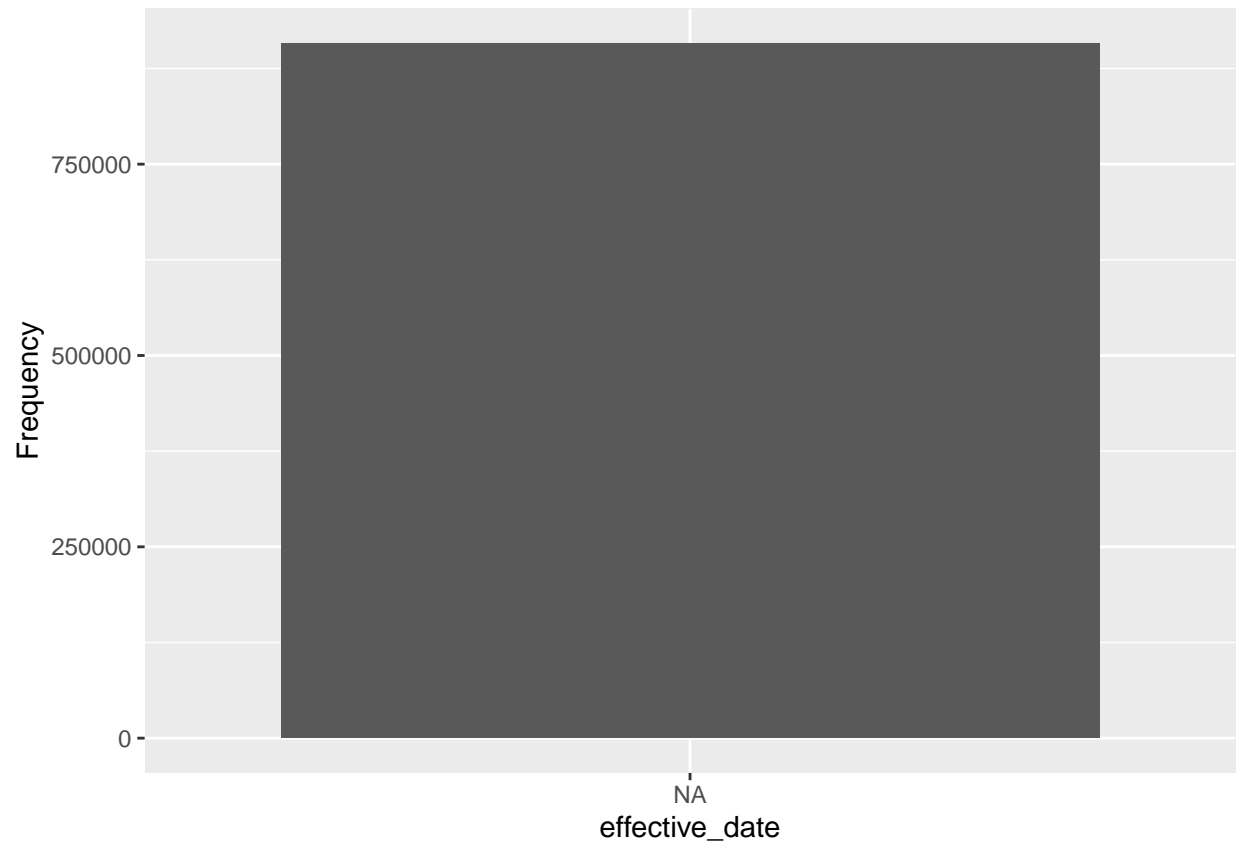
```
[1] -----  
[1] Variable: pell_grant_received, type: logical, requested: no  
[1] Values (1 unique): NA  
[1] Missing: 100%
```



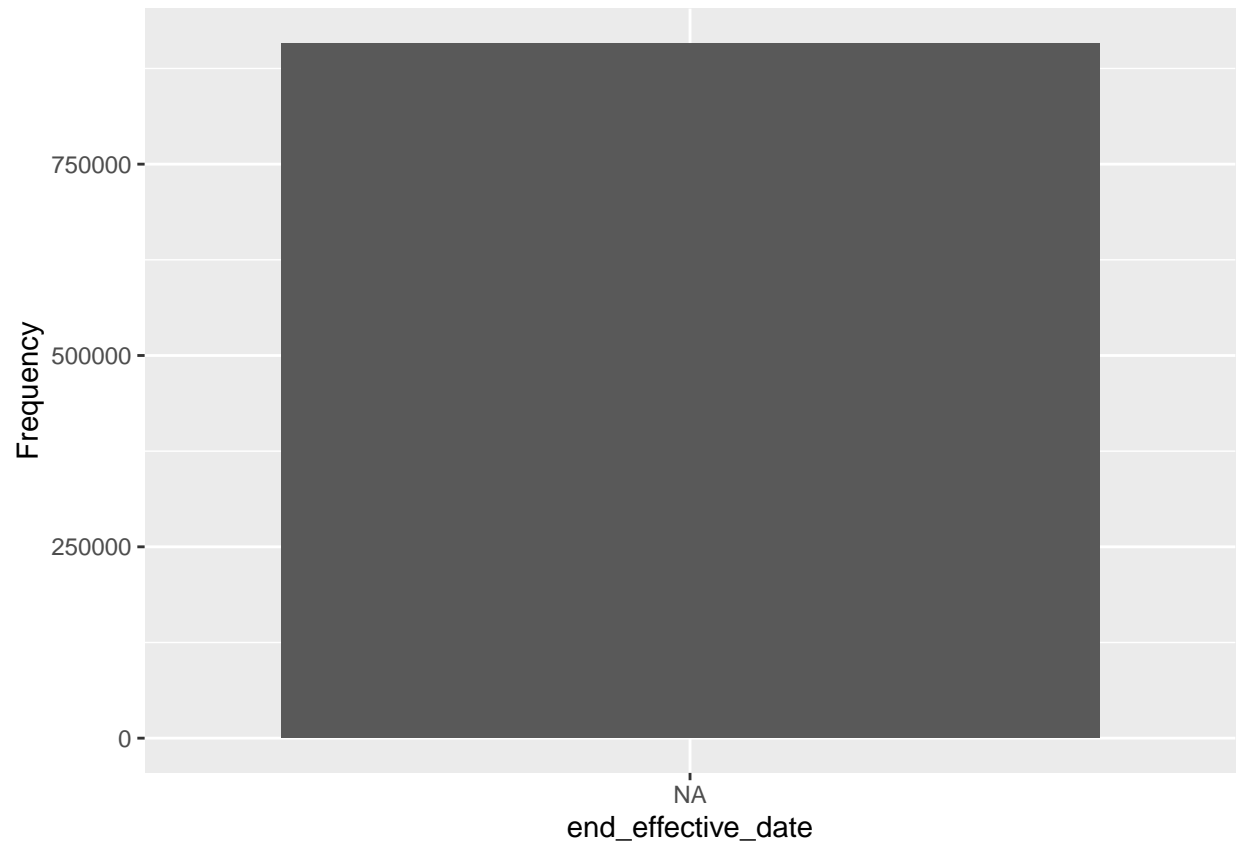
```
[1] -----  
[1] Variable: md5, type: logical, requested: no  
[1] Values (1 unique): NA  
[1] Missing: 100%
```



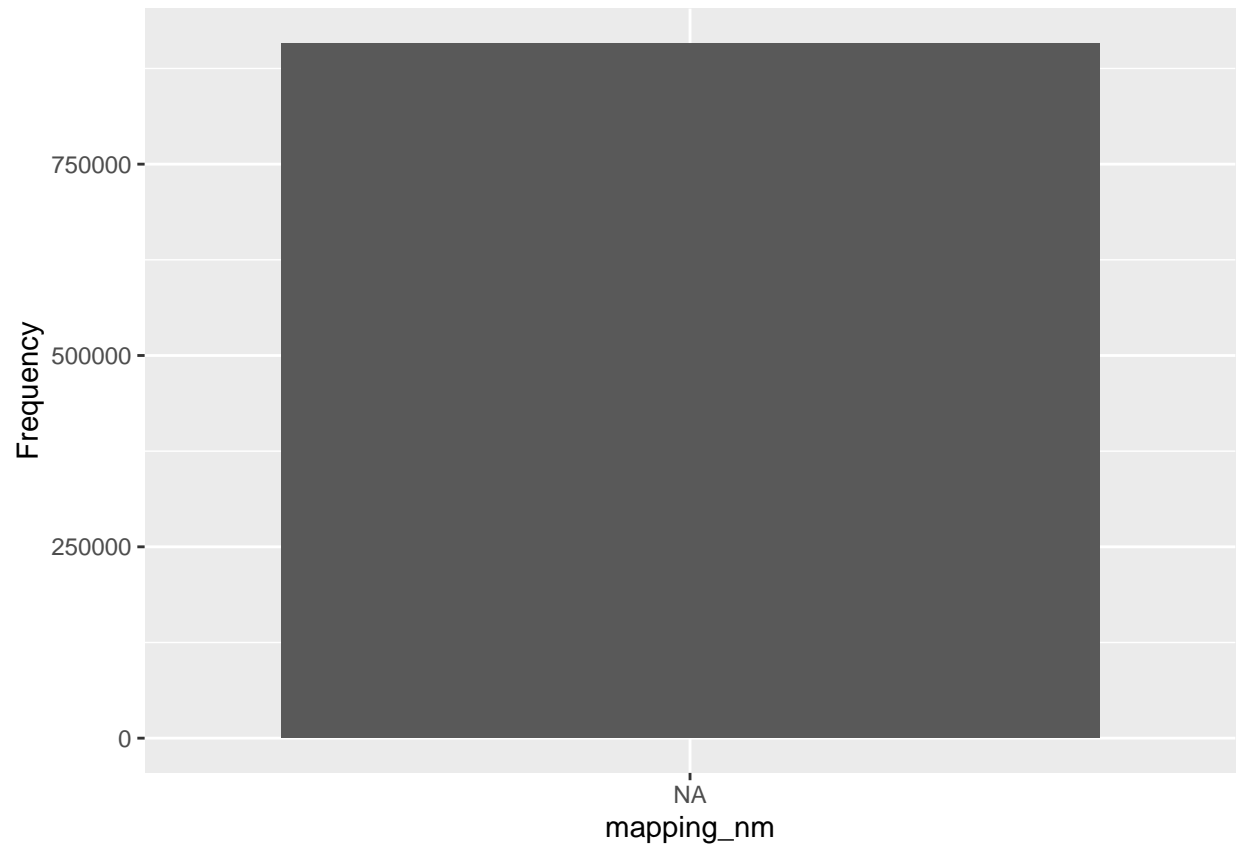
```
[1] -----  
[1] Variable: effective_date, type: logical, requested: no  
[1] Values (1 unique): NA  
[1] Missing: 100%
```



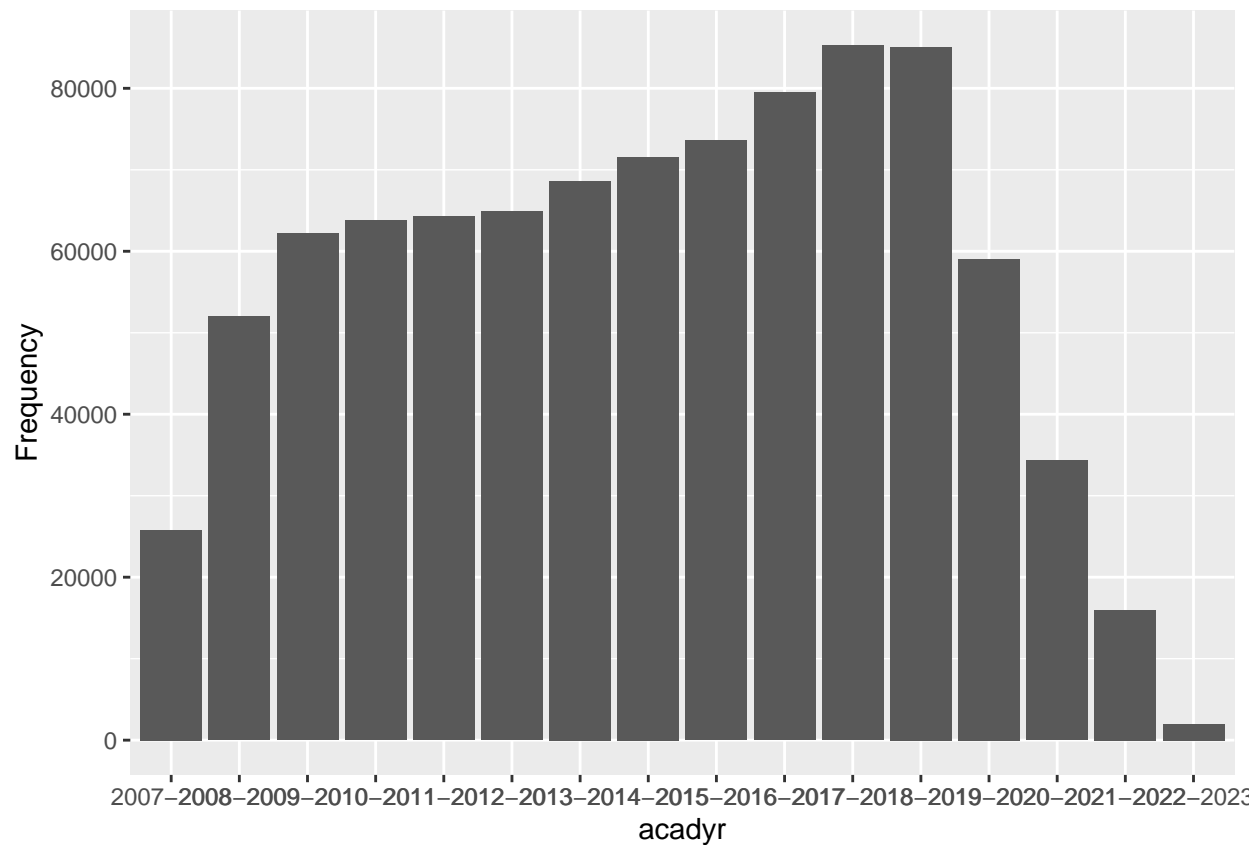
```
[1] -----  
[1] Variable: end_effective_date, type: logical, requested: no  
[1] Values (1 unique): NA  
[1] Missing: 100%
```



```
[1] -----  
[1] Variable: mapping_nm, type: logical, requested: no  
[1] Values (1 unique): NA  
[1] Missing: 100%
```

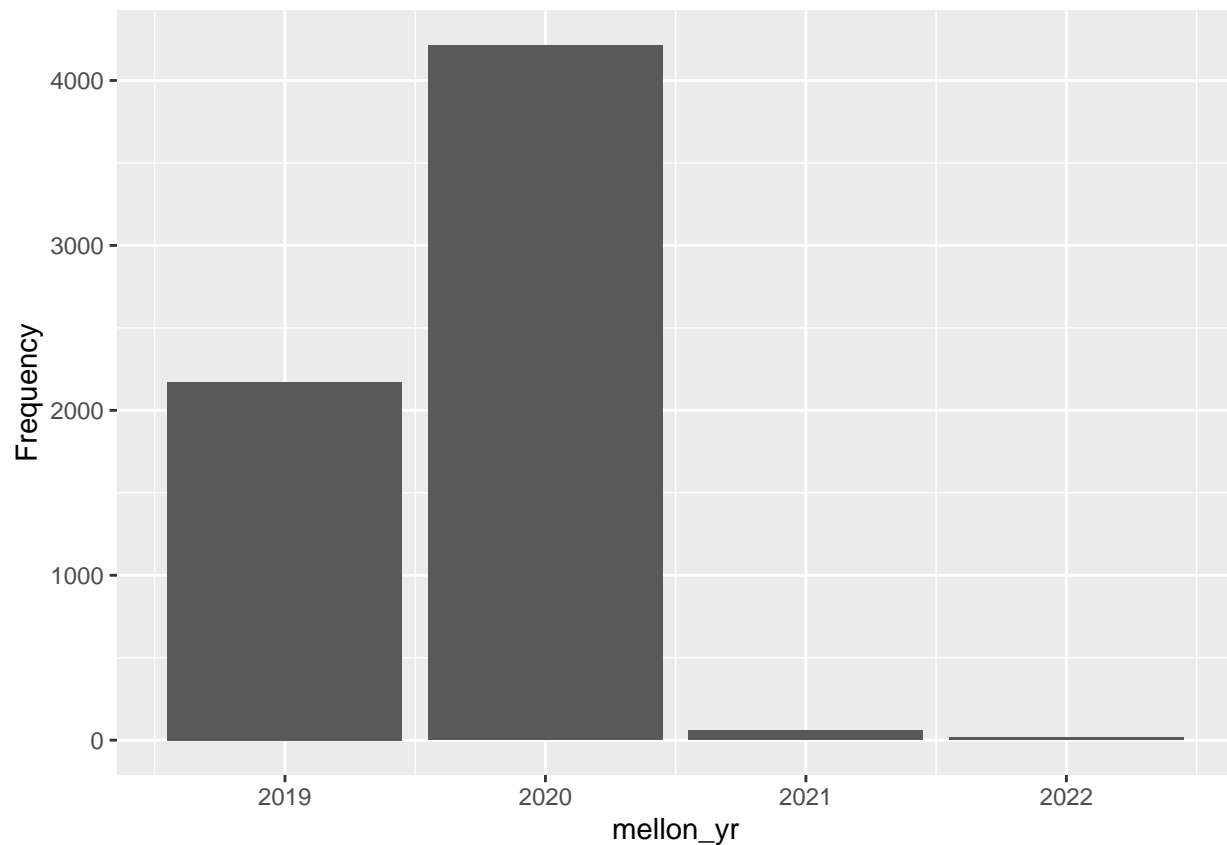


```
[1] -----  
[1] Variable: acadyr, type: character, requested: no  
[1] Values (16 unique): 2018-2019, 2019-2020, 2021-2022, 2020-2021, 2016-2017, ...  
[1] Missing: 0%
```

```
[1] -----
[1] Variable: mellon_yr, type: numeric, requested: no
[1] Values (5 unique): NA, 2020, 2019, 2022, 2021
[1] Missing: 99.3%
```

Warning: Removed 901371 rows containing non-finite values ('stat_count()').



```
[1] -----
```

These are the predictors that remain when filtering for less than 50% missing, our blacklist and variables that need to be somehow transformed first:

```
[1] "group_a"                "group_b"
[3] "term_part_code"         "term_part_desc"
[5] "citizenship"            "levelcd"
[7] "pascd"                  "active_student"
[9] "new_student"            "full_time"
[11] "year_study"             "honors"
[13] "esl_program"            "probation_status"
[15] "current_units_completed_transfer" "major_subcampus_1"
[17] "major_stem_1"
```

```
write.csv2(term_vars, 'term_vars.csv')
```