

4.01-jk-eda-report

November 17, 2025

```
[1]: # Ensure most recent (and not cached) source code is used
%load_ext autoreload
%autoreload 2
```

```
[2]: # Import libraries
import json

import matplotlib.pyplot as plt
plt.style.use("seaborn-v0_8-darkgrid")
plt.style.use("ggplot")

plt.rcParams['text.usetex'] = True
plt.rcParams['font.family'] = 'serif'
%matplotlib inline
import pandas as pd
pd.options.display.max_columns = None
import scipy
import seaborn as sns
```

```
[3]: # Import functions from other python files in the project
from pdga_score_scraper import *
from pdga_score_scraper.config import PROCESSED_DATA_DIR
```

```
2025-11-17 13:48:39.808 | INFO      |
pdga_score_scraper.config:<module>:11 - PROJ_ROOT
path is: /home/jon/repos/pdga-score-scraper
```

```
[4]: event_id = 89433
```

```
[5]: df = pd.read_json(PROCESSED_DATA_DIR / f"{event_id}-processed.json")

with open(PROCESSED_DATA_DIR / f"{event_id}-metadata-processed.json", 'r') as f:
    metadata = json.load(f)
```

```
[6]: df.head()
```

```
[6]: first_name last_name      city state country nationality pdga_num \
0    Brandon    Martin    Mohnton    PA     US     None    199207
1    Brandon    Martin    Mohnton    PA     US     None    199207
2      Mike    Carman  Ottsville    PA     US     None    48559
3      Mike    Carman  Ottsville    PA     US     None    48559
4    Andrew     Fish   Baltimore   MD     US     None    58320

    player_rating  rating_effective_date  player_class  result_id \
0            989        17604000000000  Professional  212159772
1            989        17604000000000  Professional  212159772
2            995        17604000000000  Professional  212122708
3            995        17604000000000  Professional  212122708
4           1009        17604000000000  Professional  211964506

    tournament_division  tournament_place  tied_finish formatted_place \
0                MPO                  1       False      1st
1                MPO                  1       False      1st
2                MPO                  2       False     2nd
3                MPO                  2       False     2nd
4                MPO                  4       False     4th

    shot_total  tournament_total_to_par  prize_usd  tournament_dnf \
0          125                  -11      647      False
1          125                  -11      647      False
2          127                   -9      431      False
3          127                   -9      431      False
4          132                   -4      250      False

    tournament_average_rating  tournament_rating_delta  score_id  round \
0                 1022                      33  26391524      1
1                 1022                      33  26419462      2
2                 1016                      21  26391909      1
3                 1016                      21  26419460      2
4                  999                     -10  26391664      1

    tee_time              course_name layout_name  par  round_score \
0 2025-11-17 08:48:01  Nockamixon State Park    HooDoo  68       63
1 2025-11-17 01:40:00  Nockamixon State Park    HooDoo  68       62
2 2025-11-17 08:43:01  Nockamixon State Park    HooDoo  68       64
3 2025-11-17 01:40:00  Nockamixon State Park    HooDoo  68       63
4 2025-11-17 08:44:01  Nockamixon State Park    HooDoo  68       64

    score_to_par  round_dnf  num_birdies  num_bogeys  round_rating  num_holes \
0            -5    False          8            3         1019          18
1            -6    False          7            1         1024          18
2            -4    False          7            3         1013          18
3            -5    False          6            1         1018          18
```

	4	-4	False	6	2	1013	18		
0	layout_length	Hole 1	Hole 2	Hole 3	Hole 4	Hole 5	Hole 6	Hole 7	\
1	9289	5	5	2	3	3	6	3	
2	9289	6	3	2	3	3	4	3	
3	9289	5	3	2	3	3	4	4	
4	9289	5	3	3	3	3	4	3	
	Hole 8	Hole 9	Hole 10	Hole 11	Hole 12	Hole 13	Hole 14	Hole 15	\
0	3	3	3	3	3	4	5	2	
1	4	3	4	3	3	4	4	3	
2	4	4	4	5	2	4	6	2	
3	4	4	3	3	3	4	6	3	
4	4	4	4	3	3	4	5	4	
	Hole 16	Hole 17	Hole 18						
0	3	3	4						
1	3	4	3						
2	2	3	4						
3	2	4	3						
4	2	3	3						

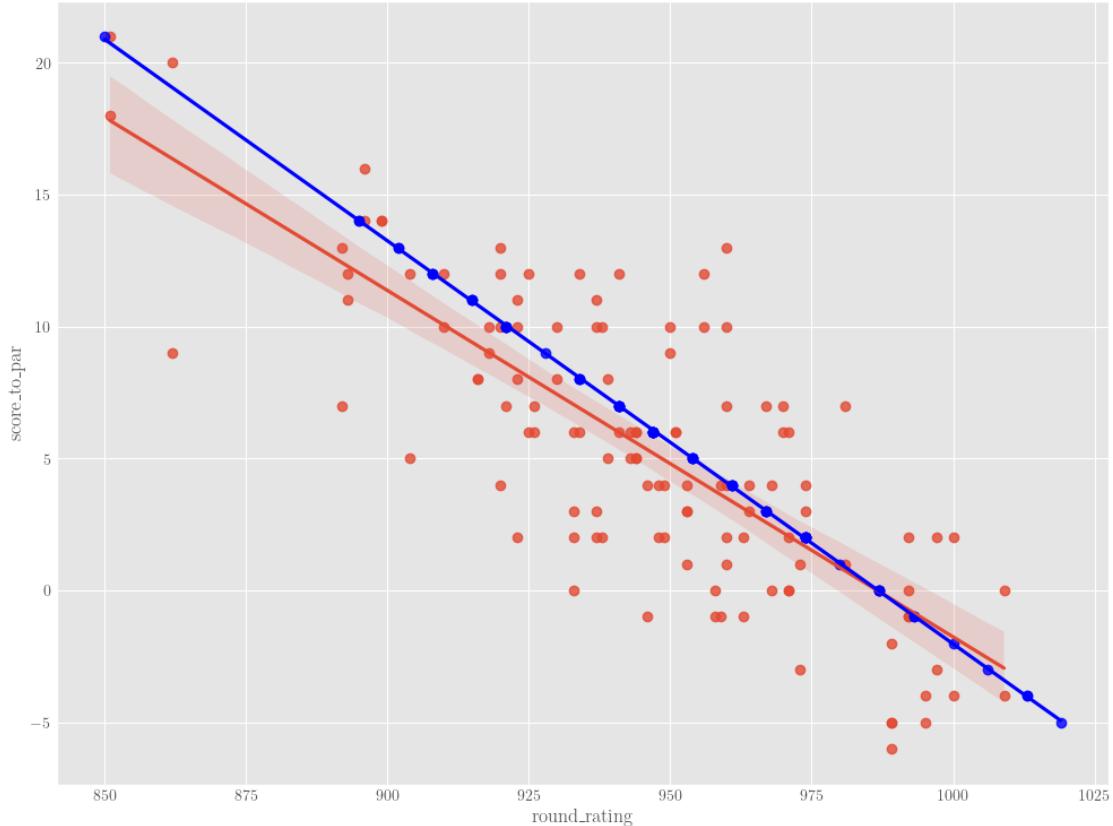
```
[7]: fig, axs = plt.subplots(1, 1, figsize = (12, 9))

sns.regplot(data=df, x="player_rating", y = "score_to_par")
#ax.set_title("Monkies")
#ax.set_xlabel("PDGA player rating")
#ax.set_ylabel("Round score relative course par")

sns.regplot(data=df[df['round'] == 1], x="round_rating", y = "score_to_par", color="blue")

#ax.legend()

plt.show()
```

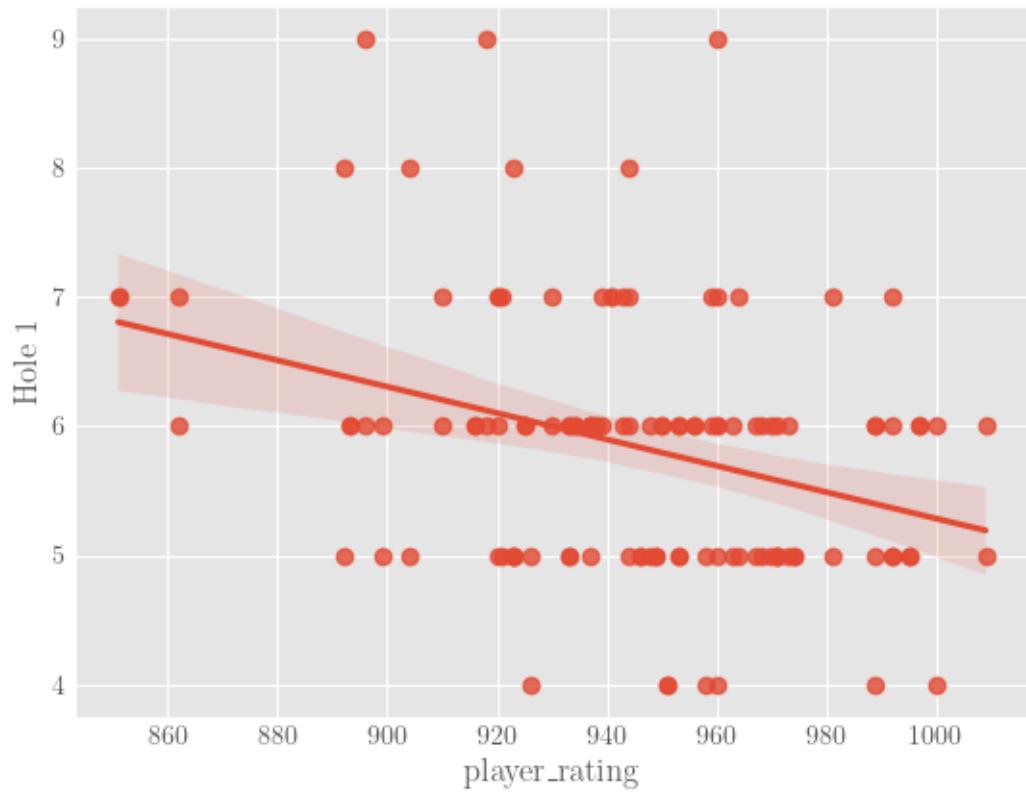


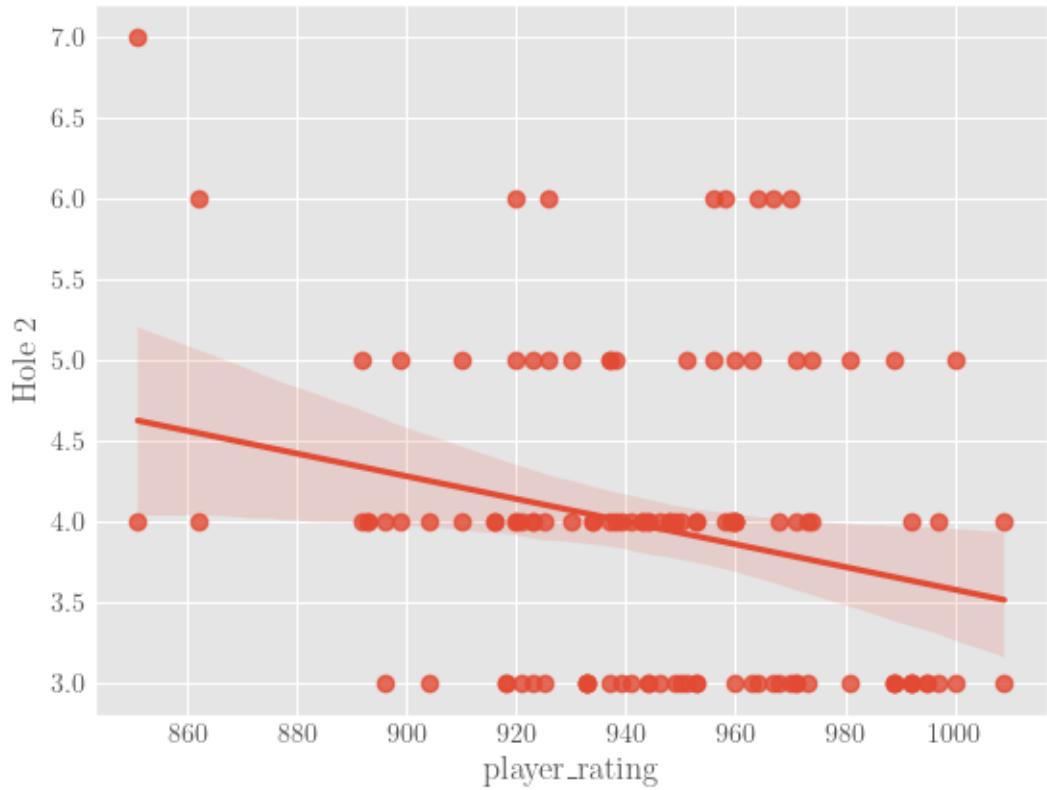
0.1 Regressing individual hole scores against player rating

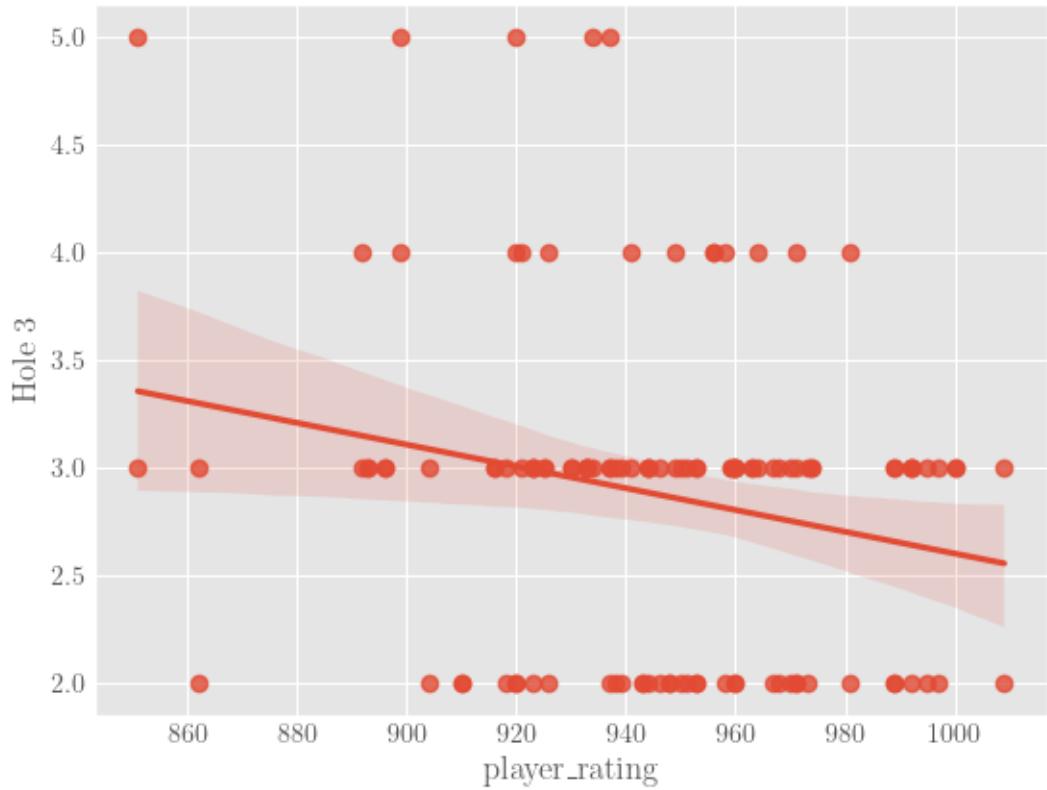
Does hole par inherently affect regression slope? Guessing yes. Maybe compare only equally-parred holes to evaluate differentiation.

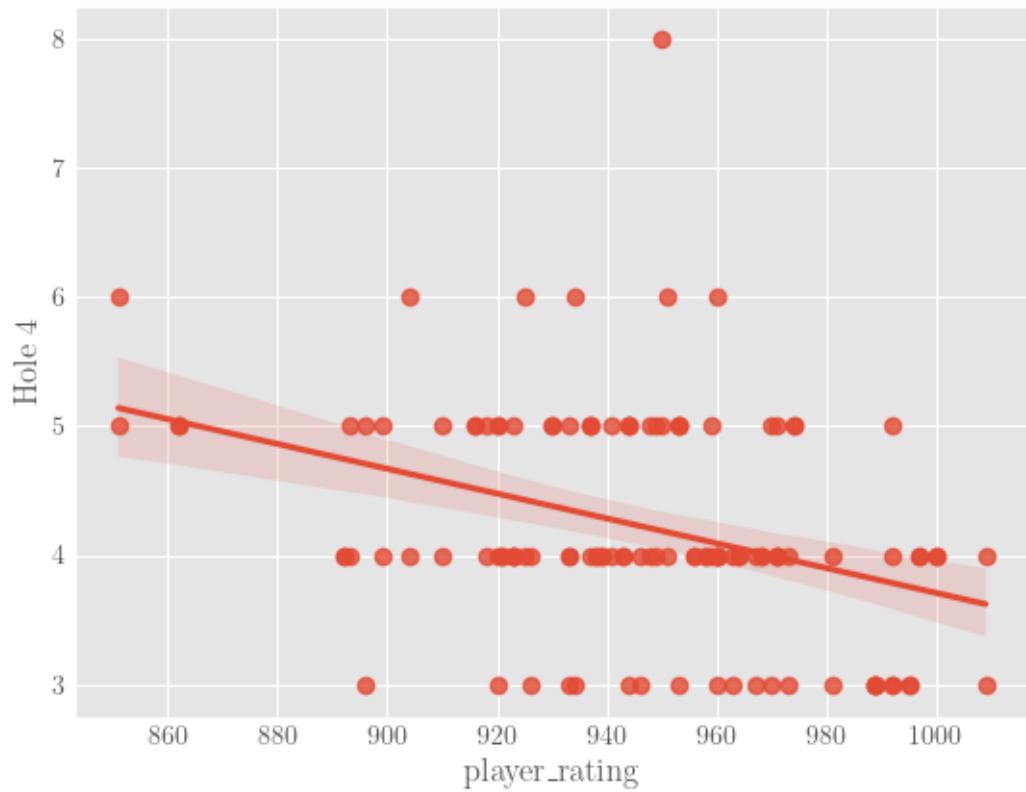
```
[8]: slopes = []
for i in range(1, 19):
    p = sns.regplot(data=df, x="player_rating", y = f"Hole {i}")

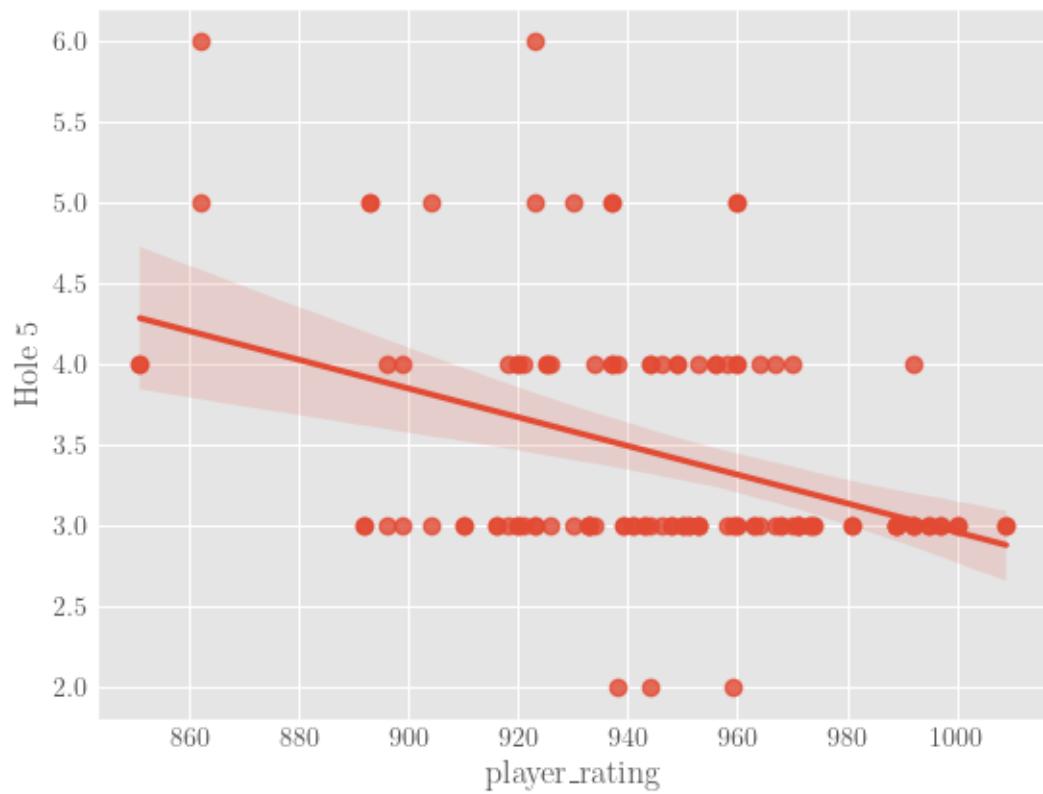
    slope, intercept, r, p, sterr = scipy.stats.linregress(x=p.get_lines()[0].
    ↪get_xdata(),
    ↪get_ydata())
    slopes[i] = slope
plt.show()
```

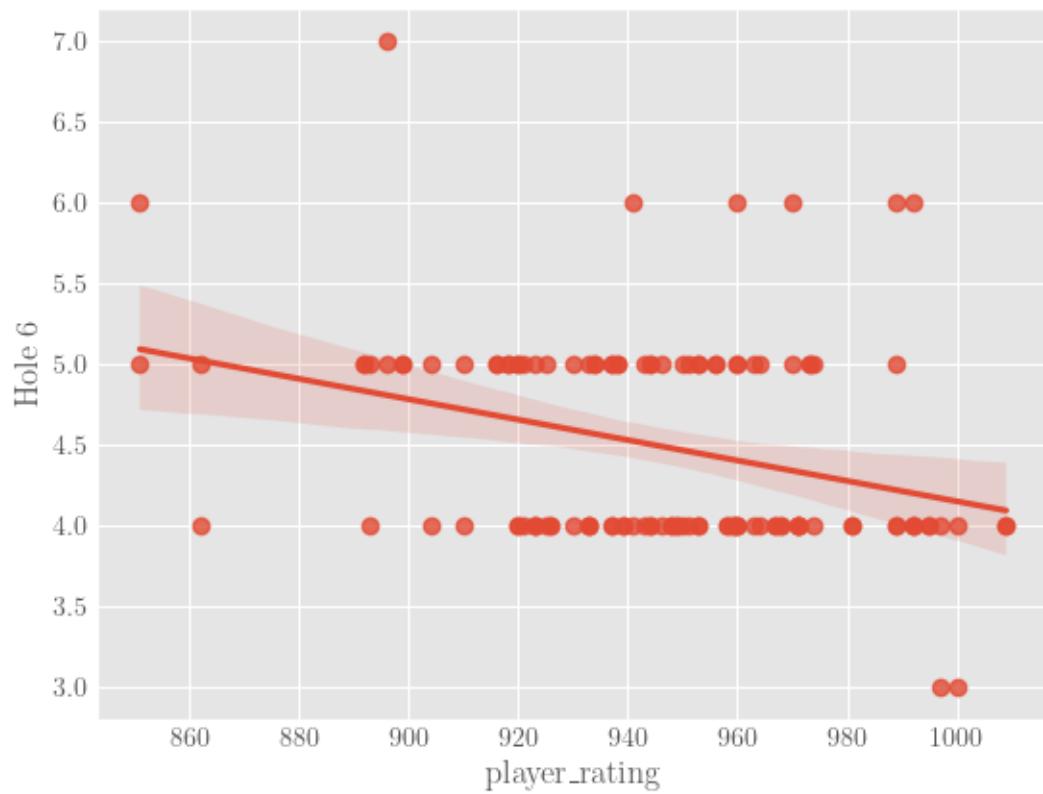


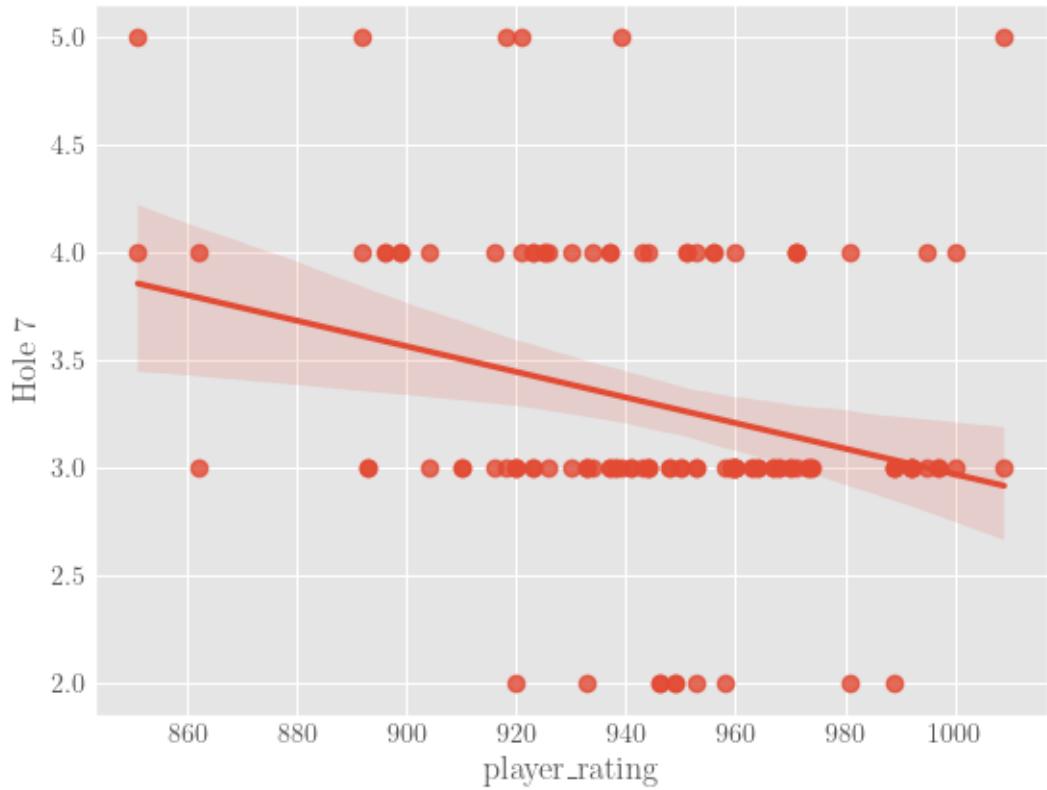


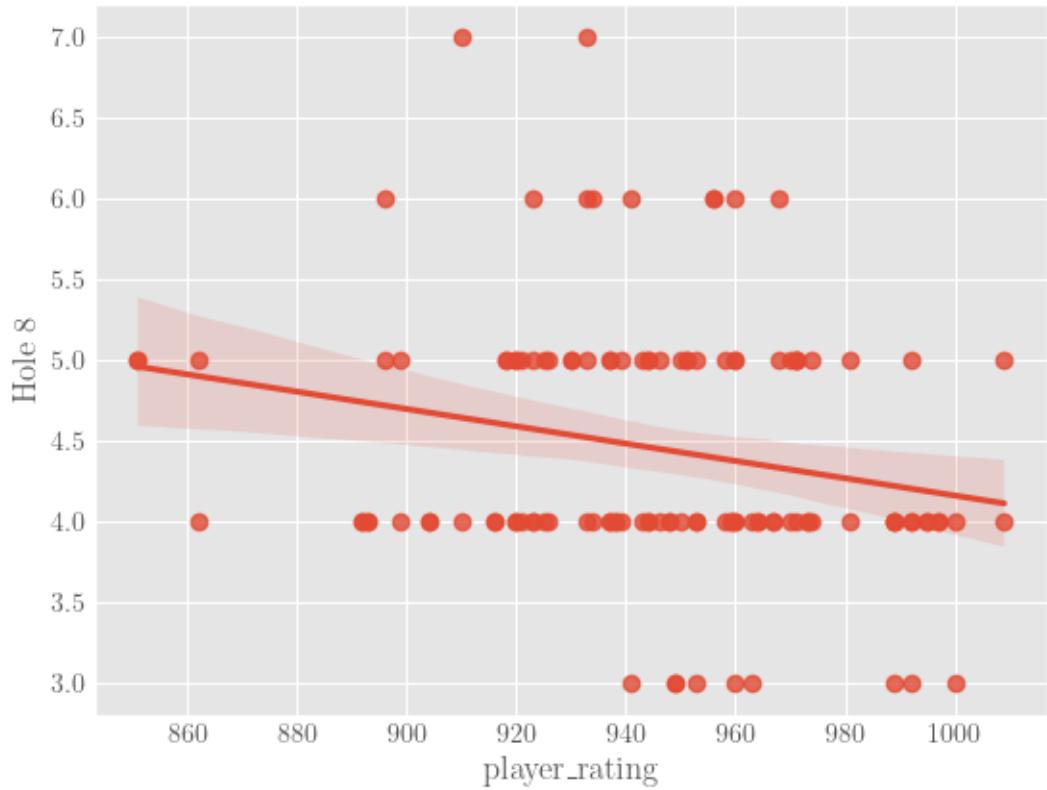


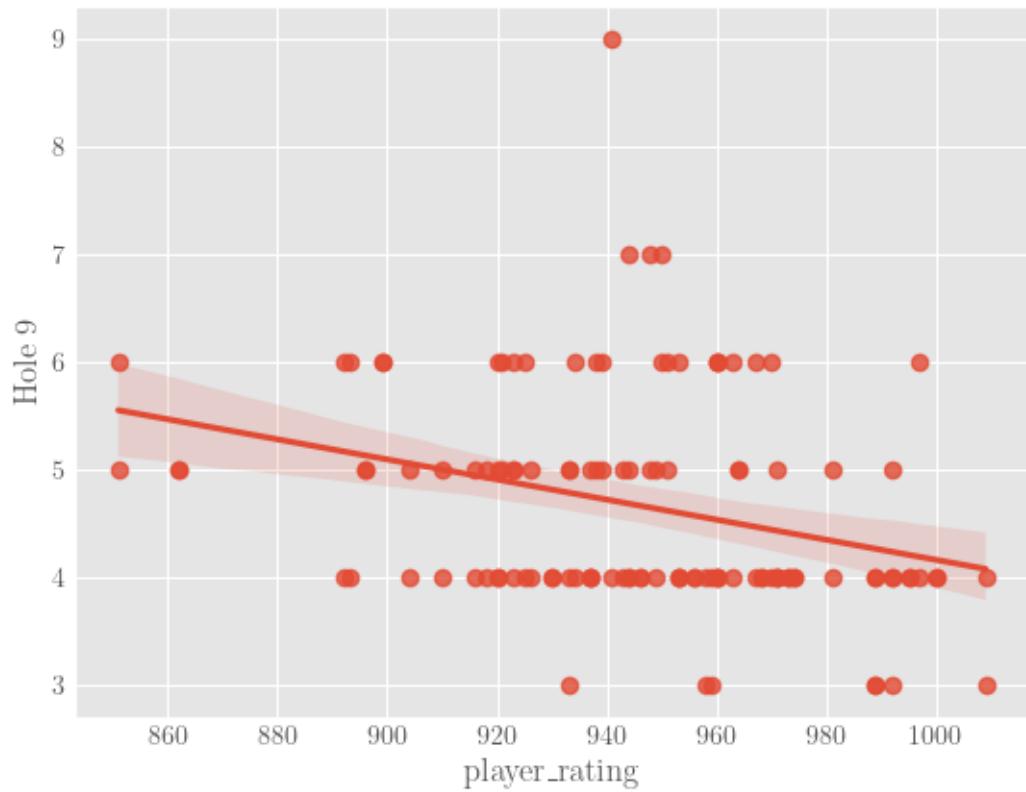


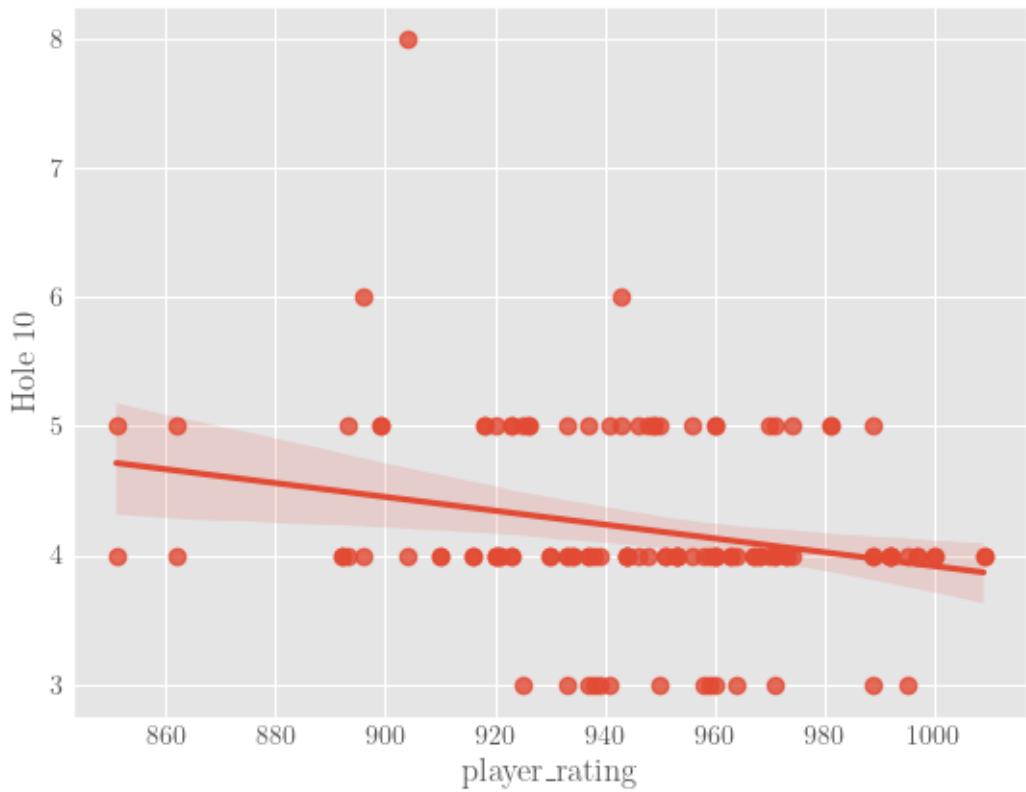


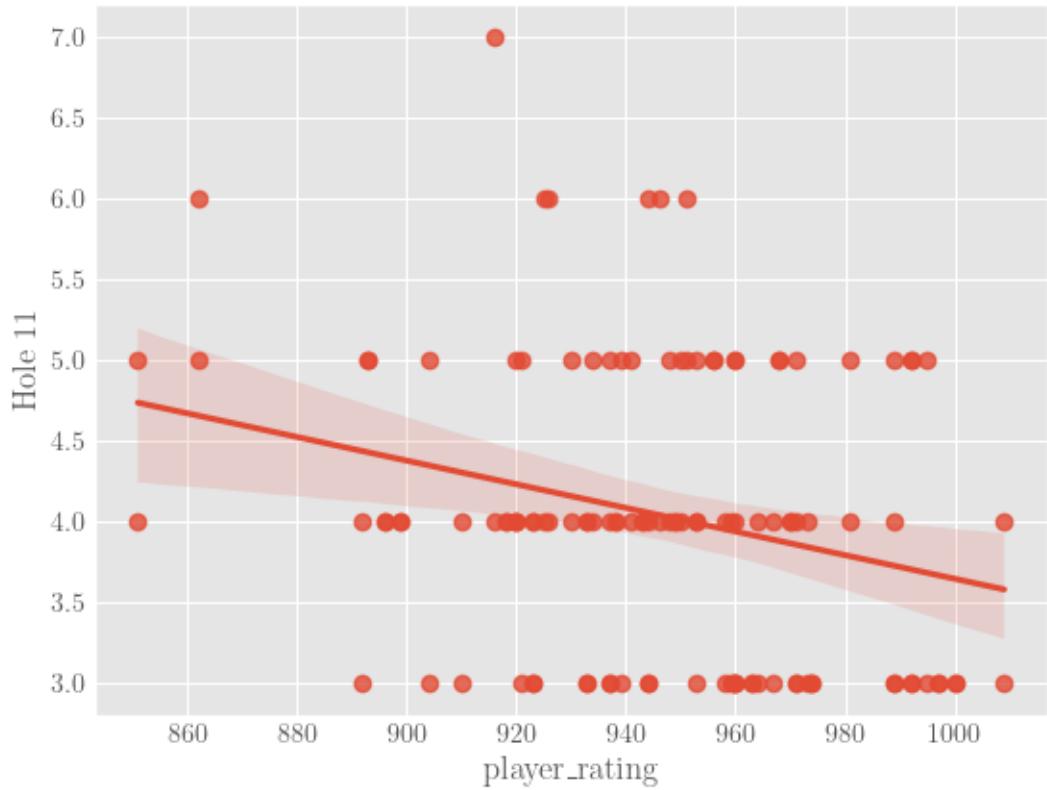


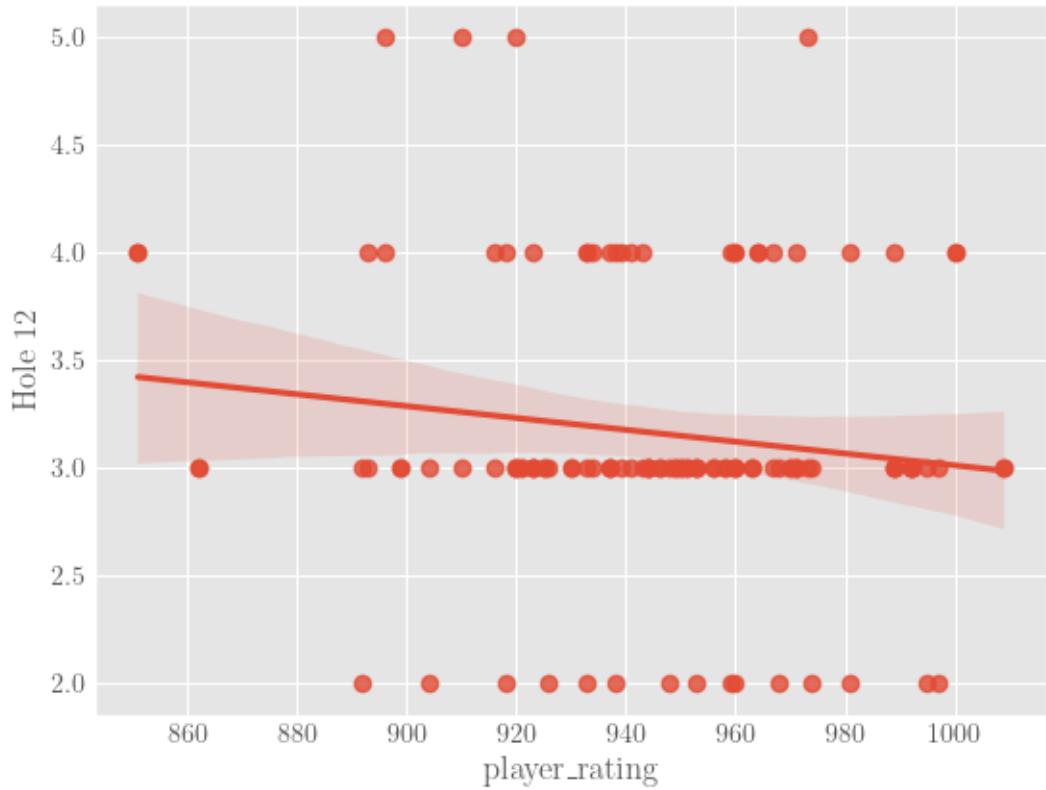


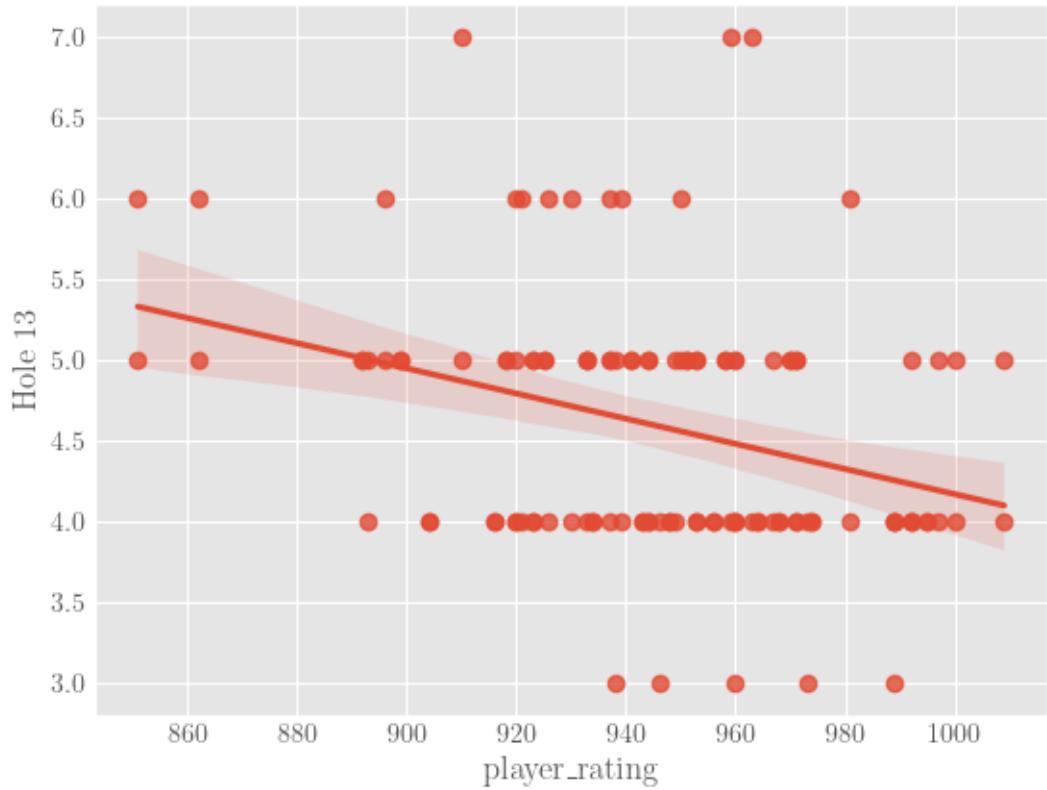


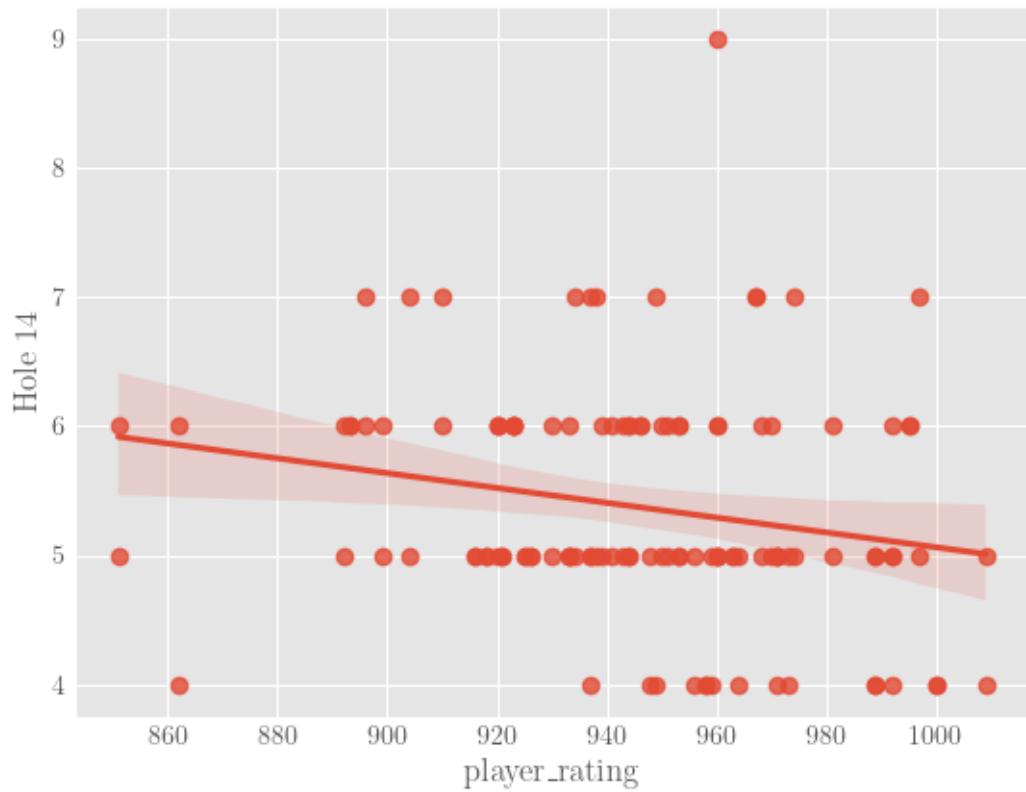


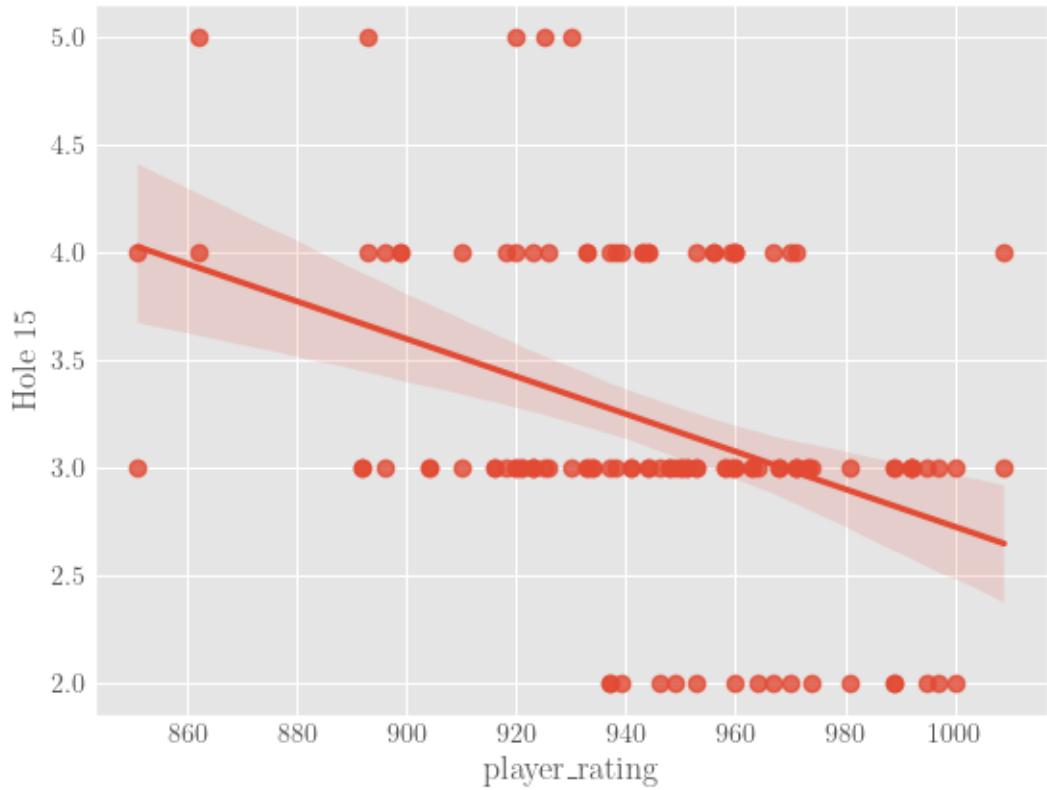


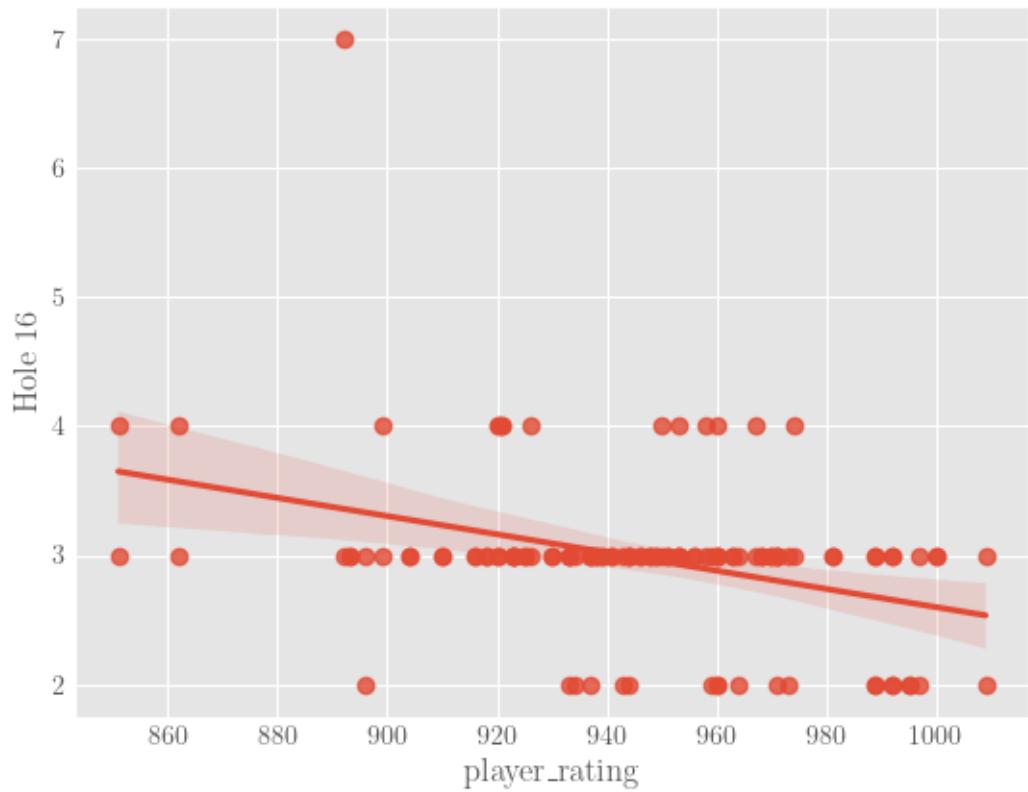


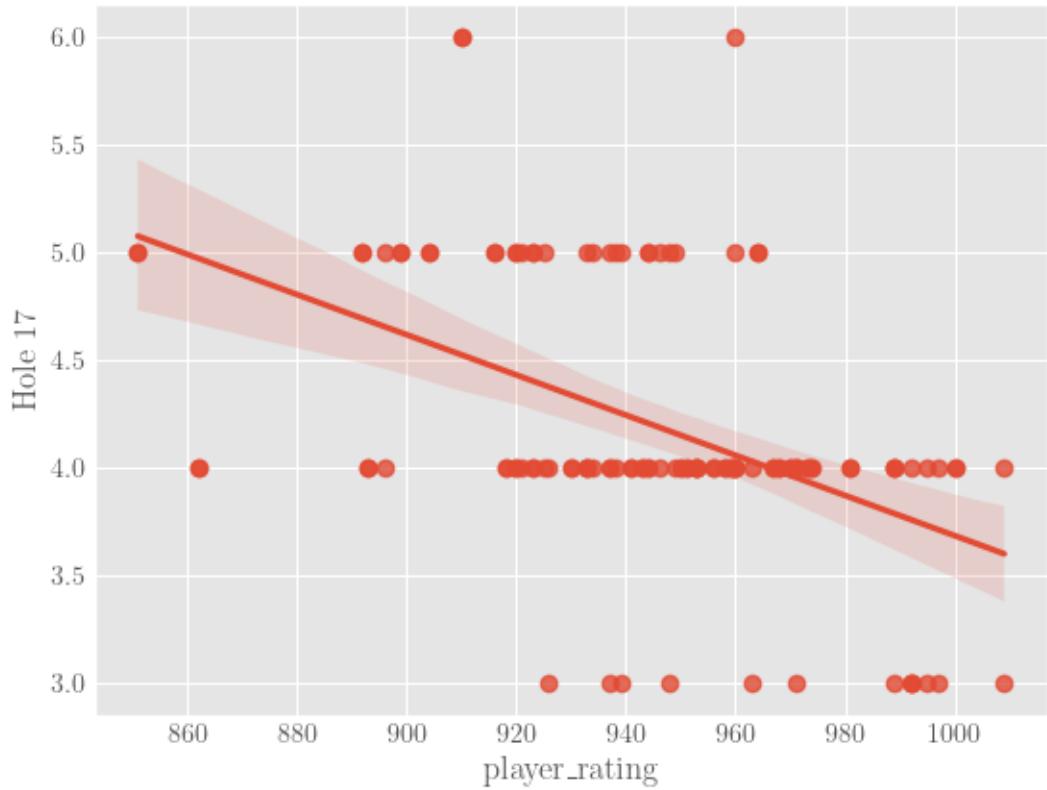


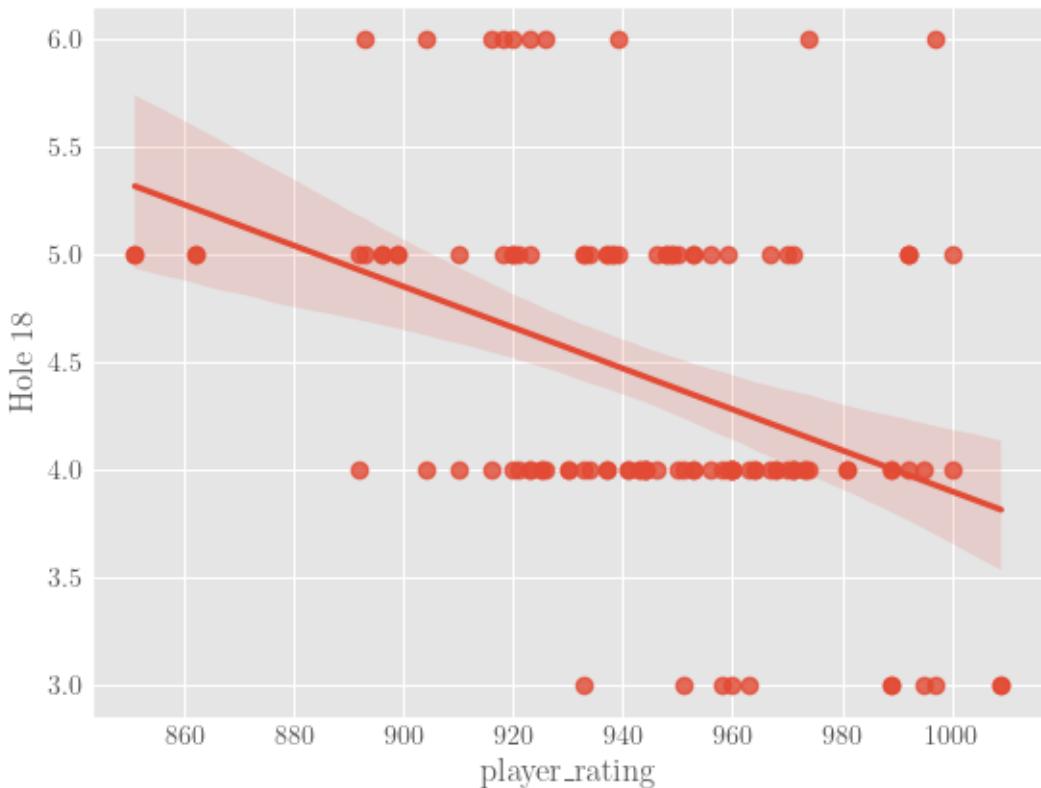




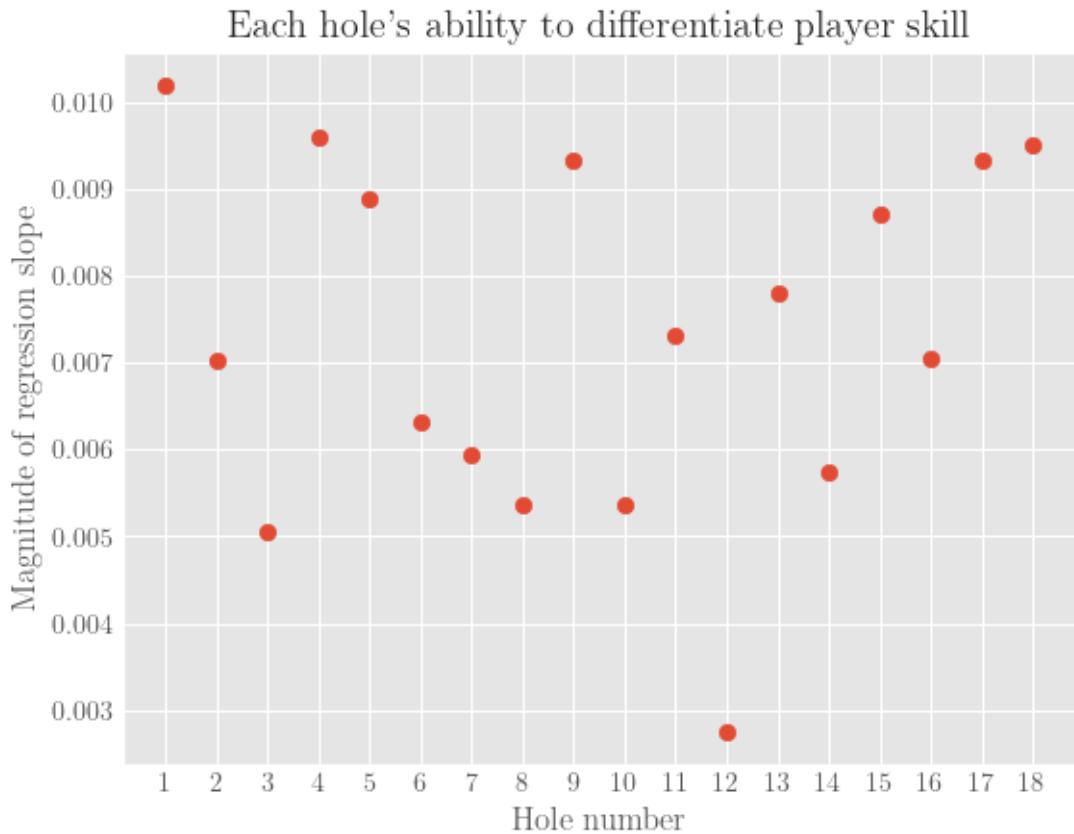








```
[9]: plt.scatter(slopes.keys(), [-1 * x for x in slopes.values()])
plt.xticks(range(1, 19))
plt.xlabel("Hole number")
plt.ylabel("Magnitude of regression slope")
plt.title("Each hole's ability to differentiate player skill")
plt.show()
```



```
[10]: s = {k: round(v, 4) for k, v in sorted(slopes.items(), key=lambda item: item[1])}
s.keys()
```

```
[10]: dict_keys([1, 4, 18, 17, 9, 5, 15, 13, 11, 16, 2, 6, 7, 14, 8, 10, 3, 12])
```

0.2 Metadata

```
[11]: metadata['tournament_info']
```

```
[11]: {'DateRange': 'Saturday, Oct 18, 2025',
'EndDate': '2025-10-18',
'Country': 'United States',
'Location': 'Quakertown, Pennsylvania',
'LocationShort': 'Quakertown, PA',
'Divisions': [{DivisionID: 100,
'Division': 'MPO',
'DivisionName': 'Mixed Pro Open',
'Players': 31,
'IsPro': True,
```

```

'ShortName': 'Mixed Pro Open',
'AbbreviatedName': 'Mixed Pro Open',
'LatestRound': 2},
{'DivisionID': 111,
'Division': 'MP40',
'DivisionName': 'Mixed Pro 40+',
'Players': 4,
'IsPro': True,
'ShortName': 'Mixed Pro 40+',
'AbbreviatedName': 'Mixed Pro 40+',
'LatestRound': 2},
{'DivisionID': 121,
'Division': 'MP50',
'DivisionName': 'Mixed Pro 50+',
'Players': 5,
'IsPro': True,
'ShortName': 'Mixed Pro 50+',
'AbbreviatedName': 'Mixed Pro 50+',
'LatestRound': 2},
{'DivisionID': 400,
'Division': 'MA1',
'DivisionName': 'Mixed Amateur 1',
'Players': 20,
'IsPro': False,
'ShortName': 'Mixed Amateur 1',
'AbbreviatedName': 'Mixed Am 1',
'LatestRound': 2}],
'Finals': 'no',
'HighestCompletedRound': 2,
'LatestRound': 2,
'FinalRound': 2,
'Layouts': [ {'LayoutID': 721578,
'CourseID': 25812,
'CourseName': 'Nockamixon State Park',
'Name': 'HooDoo',
'Holes': 18,
'Par': 68,
'Length': 9289,
'Units': 'Feet',
'Accuracy': 'M',
'Details': [ {'Hole': 'H1', 'Label': '1', 'Par': 5, 'Length': 845},
{'Hole': 'H2', 'Label': '2', 'Par': 4, 'Length': 447},
{'Hole': 'H3', 'Label': '3', 'Par': 3, 'Length': 220},
{'Hole': 'H4', 'Label': '4', 'Par': 4, 'Length': 644},
{'Hole': 'H5', 'Label': '5', 'Par': 3, 'Length': 304},
{'Hole': 'H6', 'Label': '6', 'Par': 4, 'Length': 626},
{'Hole': 'H7', 'Label': '7', 'Par': 3, 'Length': 390}]}

```

```

{'Hole': 'H8', 'Label': '8', 'Par': 4, 'Length': 570},
{'Hole': 'H9', 'Label': '9', 'Par': 4, 'Length': 541},
{'Hole': 'H10', 'Label': '10', 'Par': 4, 'Length': 565},
{'Hole': 'H11', 'Label': '11', 'Par': 4, 'Length': 404},
{'Hole': 'H12', 'Label': '12', 'Par': 3, 'Length': 317},
{'Hole': 'H13', 'Label': '13', 'Par': 4, 'Length': 617},
{'Hole': 'H14', 'Label': '14', 'Par': 5, 'Length': 769},
{'Hole': 'H15', 'Label': '15', 'Par': 3, 'Length': 358},
{'Hole': 'H16', 'Label': '16', 'Par': 3, 'Length': 377},
{'Hole': 'H17', 'Label': '17', 'Par': 4, 'Length': 542},
{'Hole': 'H18', 'Label': '18', 'Par': 4, 'Length': 753}]}],
'Name': 'The Hoodoo Presented by OTB East and Discraft - MPO, FPO, MP40, FP40,
MP50, FP50, MA1',
'NameHtml': 'The Hoodoo Presented by OTB East and Discraft - MPO, FPO, MP40,
FP40, MP50, FP50, MA1',
'MultiLineName': {'pre': '',
'main': 'The Hoodoo',
'post': 'Presented by OTB East and Discraft - MPO, FPO, MP40, FP40, MP50,
FP50, MA1'},
'MultiLineNameHtml': '<span class="event-name-main">The Hoodoo</span> <span
class="event-name-post">Presented by OTB East and Discraft - MPO, FPO, MP40,
FP40, MP50, FP50, MA1</span>',
'RawTier': 'B',
'Rounds': 2,
'RoundsList': {'1': {'Number': 1,
'Label': 'Round 1',
'LabelAbbreviated': 'Rd 1',
'Date': '2025-10-18',
'DateAbbreviated': 'Oct 18',
>ShowDate': False},
'2': {'Number': 2,
'Label': 'Round 2',
'LabelAbbreviated': 'Rd 2',
'Date': '2025-10-18',
'DateAbbreviated': 'Oct 18',
>ShowDate': False}},
'Semis': 'no',
'SimpleName': 'The Hoodoo',
'StartDate': '2025-10-18',
'TDName': 'Steve Braud',
'TDPDGANum': 16027,
'Tier': 'B',
'FormattedTier': 'B-Tier',
'FormattedLongTier': 'B-Tier',
'TierPro': 'B',
'TierAm': 'B',
'TierX': False,

```

```
'TotalPlayers': 60,  
'TournamentId': '89433',  
'International': False,  
'FullStats': True,  
'TimeZone': 'America/New_York',  
'AdditionalEventInfo': None,  
'EventVideo': None,  
'ScoringFormat': 'S'}
```

[]: