# Exploring Consumer Tastes and Preferences in Indianapolis Golf Courses

Alex Heck

Unstructured Data Analytics

# Data Source





- GolfPass.com – direct affiliate of NBC Sports and Golf Channel

- Tee time bookings, video instruction, travel advice, courses, and daily tips

- I scraped comment review data from the "Travel" tab, and the *Indianapolis Golf Guide* section.

- Comments were embedded in hyperlinks. First scraped unique course IDs then iterated through 100+ URLs and approx. 10-16 pages to extract the comments. *(code took around 55 minutes to run)*

# Goals

Generate my own unique rankings based on sentiment scores and uncover what golfers truly care about in a course.

# The Numbers

**103** Golf Courses in Indianapolis area

**6458** user reviews

**I observed that some courses had very few reviews**
Filtered to only analyze courses that had <u>more than 20 reviews</u>:

**45** Indianapolis Golf Courses

**6330** user reviews

# Sentiment Analysis

◎ I used **Vader** to perform sentiment analysis over each review

◎ Then I grouped the reviews by course

◎ Finally combined the reviews by calculating the mean compound sentiment score for each course (descending order)

# Top 5 Courses in Indianapolis (average sentiment)



**Buffer Park**
0.8016

**Deer Creek**
0.7572

**Bear Slide**
0.7115

**Arrowhead**
0.7102

**Eagle Pines**
0.7038

| | course_name | average_sentiment_score |
|---|---|---|
| 0 | Buffer Park Golf Course | 0.801620 |
| 1 | Deer Creek Golf Club | 0.757170 |
| 2 | Bear Slide Golf Club | 0.711486 |
| 3 | Arrowhead Golf Course | 0.710173 |
| 4 | Eagle Pines Golf Club | 0.703826 |
| 5 | Valle Vista Country Club | 0.700227 |
| 6 | Plum Creek Golf Club | 0.675907 |
| 7 | Brickyard Crossing | 0.674500 |
| 8 | Fox Prairie Golf Club - The Central/East Course | 0.669152 |
| 9 | Pebble Brook Golf Club - North Course | 0.665901 |
| 10 | South Grove Golf Course | 0.654762 |
| 11 | Golf Club of Indiana | 0.643226 |
| 12 | The Legends Golf Club - Creek/Middle Course | 0.641496 |
| 13 | Riverside Golf Course | 0.633872 |
| 14 | Eagle Creek Golf Club - The Pines Course | 0.624124 |
| 15 | The Trophy Club | 0.622270 |
| 16 | Foxcliff Golf Club | 0.618761 |
| 17 | Winding River Golf Club | 0.616858 |
| 18 | Pebble Brook Golf Club - South Course | 0.611749 |
| 19 | Prairie View Golf Club | 0.607346 |
| 20 | Smock Golf Course | 0.606006 |
| 21 | Hickory Stick Golf Club | 0.605934 |
| 22 | The Fort Golf Resort | 0.603110 |
| 23 | River Glen Country Club | 0.591196 |
| 24 | Southern Dunes Golf Course | 0.583911 |
| 25 | West Chase Golf Club | 0.578456 |
| 26 | The Lakes/The Ridge at Ironwood Golf Club | 0.577727 |
| 27 | Brookshire Golf Course | 0.576283 |
| 28 | Wood Wind Golf Club | 0.567973 |
| 29 | Twin Bridges Golf Club | 0.567201 |
| 30 | Purgatory Golf Club | 0.563029 |
| 31 | Hawk's Tail of Greenfield | 0.541243 |
| 32 | Sahm Golf Course | 0.529025 |
| 33 | Links Golf Club, The | 0.526265 |
| 34 | Coffin Golf Club | 0.516594 |
| 35 | Bluff Creek Golf Course | 0.497388 |
| 36 | The Links at Heartland Crossing | 0.475651 |
| 37 | Bear Chase Golf Club | 0.467119 |
| 38 | Eagle Creek Golf Club - The Sycamore Course | 0.466439 |
| 39 | Saddlebrook Golf Club | 0.427210 |
| 40 | Gray Eagle Golf Club | 0.424113 |
| 41 | Winding Ridge Golf Club | 0.418030 |
| 42 | Oaktree Golf Course | 0.351835 |
| 43 | Morningstar Golf Club | 0.324304 |
| 44 | Prestwick Country Club | 0.250201 |

# Topic Models

◉ Used BERT to perform topic modeling

Conducted 3 different topic models:
1. Entire Dataset
2. Reviews from my home golf course (The Legends)
3. Negative review comments (sentiment < .10)

## All Reviews/Courses

| Topic | Count | Name | Representation |
|---|---|---|---|
| -1 | 2635 | -1_boxes_fairways_ball_grass | [boxes, fairways, ball, grass, rough, fairway,... |
| 0 | 426 | 0_us_he_told_said | [us, he, told, said, minutes, twosome, behind,... |
| 1 | 328 | 1_indy_indianapolis_indiana_central | [indy, indianapolis, indiana, central, courses... |
| 2 | 242 | 2_bunkers_sand_bunker_rocks | [bunkers, sand, bunker, rocks, them, need, unp... |
| 3 | 194 | 3_here_playing_again_definitely | [here, playing, again, definitely, always, int... |
| 4 | 193 | 4_staff_friendly_helpful_always | [staff, friendly, helpful, always, act, drinks... |
| 5 | 138 | 5_golf_family_place_steve | [golf, family, place, steve, golfers, golfer, ... |
| 6 | 125 | 6_rain_wet_raining_weather | [rain, wet, raining, weather, rained, dry, sta... |
| 7 | 103 | 7_nine_front_back_neighborhood | [nine, front, back, neighborhood, different, t... |
| 8 | 88 | 8_fairways_heat_rough_rolled | [fairways, heat, rough, rolled, stress, rollin... |

## The Legends Golf Club
### (my home course)

| Topic | Count | Name | Representation |
|---|---|---|---|
| -1 | 57 | -1_to_and_the_of | [to, and, the, of, we, was, in, but, on, for] |
| 0 | 32 | 0_course_very_again_great | [course, very, again, great, was, deal, time, ... |
| 1 | 26 | 1_holes_27_all_to | [holes, 27, all, to, they, on, play, is, golfe... |
| 2 | 23 | 2_greens_in_shape_and | [greens, in, shape, and, are, fast, course, we... |
| 3 | 13 | 3_legends_the_of_year | [legends, the, of, year, is, than, great, golf... |
| 4 | 12 | 4_rain_wet_was_raining | [rain, wet, was, raining, course, it, all, sha... |

## Negative Reviews

| Topic | Count | Name | Representation |
|---|---|---|---|
| -1 | 357 | -1_boxes_tee_conditions_into | [boxes, tee, conditions, into, just, grass, re... |
| 0 | 63 | 0_bunkers_sand_rocks_unplayable | [bunkers, sand, rocks, unplayable, them, dirt,... |
| 1 | 49 | 1_nine_hole_par_shot | [nine, hole, par, shot, holes, front, dogleg, ... |
| 2 | 45 | 2_told_online_now_golfnow | [told, online, now, golfnow, booked, guy, golf... |
| 3 | 39 | 3_pay_price_42_total | [pay, price, 42, total, deal, charged, senior,... |
| 4 | 27 | 4_ranger_rangers_pace_behind | [ranger, rangers, pace, behind, pickup, monito... |
| 5 | 25 | 5_walker_cart_walkable_walkers | [walker, cart, walkable, walkers, carts, push,... |
| 6 | 22 | 6_fungus_nor_fast_smooth | [fungus, nor, fast, smooth, low, yearnever, 2p... |
| 7 | 21 | 7_aeration_speeds_fast_putting | [aeration, speeds, fast, putting, birds, disap... |
| 8 | 20 | 8_ruff_dead_brown_weeds | [ruff, dead, brown, weeds, poorly, score, blue... |

# Conclusions

- Can look at each topic model and derive insights
- I found value in looking across all the models and finding overlap/common themes

- **1**. Bunkers & sand quality   **2**. Staff/Employees   **3.** Pace of play    **4.** Greens

- Courses looking to improve ratings and enhance overall experience should focus on these areas

# Thank you

aheck3@nd.edu

# Sources:

https://www.golfpass.com/ (GolfPass Website)

https://www.golfpass.com/travel-advisor/destinations/57-indianapolis-in/#courses (GolfPass Indianapolis Courses)