Athena Softworks-Customer Data Project

Hoffmann Zhu

Background

Athena Softworks, Inc. is a video game developer and publisher. They specialize in premium role-playing games (RPGs) for PC play and have published eight games to date. Athena distributes its games exclusively using Steam, the world's largest distributor of PC games, which takes up 75% of the global market share. It is early 2020 and Athena is considering acquiring a new game title and there are three candidate games.

Candidate Games

Warrior Guild is a multiplayer battle arena game that can be played in duel, team (groups of 2, 3, 5, or 7), or guild (groups of 12) modes. Players compete as one of seven classes: Vanguard, Mage, Shapeshifter, Marksman, Healer, Rogue, and Berserker.

Seraph Guardians is an immersive single player RPG with extensive strategy and puzzle elements. Players work to discover the mythology of the Seraph through strategy, problem-solving, and combat. There is more than one way to accomplish every goal, but not all methods are equally efficient.

Evercrest may be played in single player or multiplayer modes as players work to save the titular land. Players should prepare to be met with challenges that require utmost creativity as they explore this fanciful world, encountering compelling characters and narrative.

Financial Details

In all instances, Athena would pay 5% in royalties (on gross) to the original developer and spend $7 million in fixed costs to acquire and market the game; each game would have unique additional costs to finalize development ($6 million for Evercrest, $5.5 million for Seraph Guardians, and $5 million for Warrior Guild). Budget constraints dictate that only one game may be acquired. Any game not acquired by Athena may be released by another publisher.

Steam is a digital video game distribution service which holds about 75% of the PC game global market share with approximately 18 million users. For games sold on Steam, parent company Valve takes 30% of gross sales up to the first $10 million. For all sales between $10 million and $50 million, Valve takes 25%. For every sale after the initial $50 million, Valve takes 20%. You may assume that the majority of games sales occur within the first year of release and that all games sold by Athena Softworks are sold via Steam.



Analytics

1. **<u>Decision 1: Which game should Athena pursue, if any?</u>**
    - Four action alternatives:
        - Only acquire Warrior Guild,
        - Only acquire Seraph Guardians,
        - Only acquire Evercrest
        - Do not acquire any game.

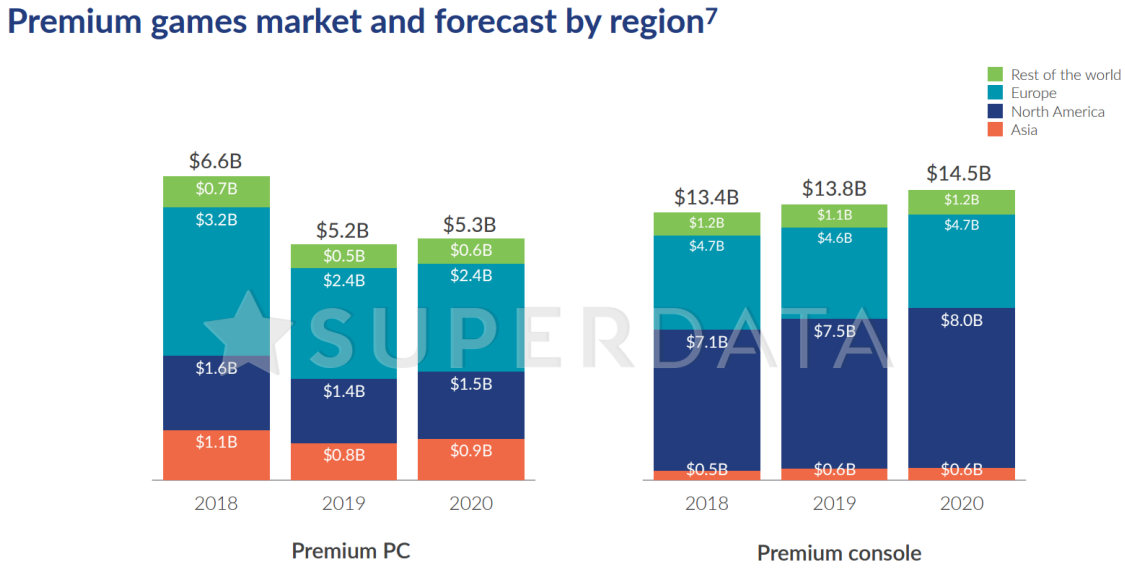
    - To evaluate the action alternatives we need action standards which are listed below:
        - **Market Potential and Target Market:** Who is our target audience? What is the potential market size for each game?
        - **Financial Projections:** What is the ROI of the acquired game? What is the Net Present Value (NPV) of the acquired game? We would need to take development costs, expected sales and the revenue-sharing model with Steam (Valve) into account.
        - **Competitive Landscape:** What will other publishers do if Athena does not acquire any game? What game is most likely going to be acquired by competitors?
        - **Strategic Fit:** Assess how well the acquired game aligns with Athena's game portfolio and strategic goals.

    <u>Decision 2: How should the game be priced?</u>
    - Again, four action alternatives:
        - **Premium Pricing:** Price the game at a premium if the acquired game is perceived as superior in quality.
        - **Competitive Pricing:** Align the price with similar games in the market.
        - **Penetration Pricing:** Set the price lower to gain market share at a rapid pace.
        - **Dynamic Pricing:** Adjust the price of the game according to demand, sales performance and promotions.
    - Action standards
        - **Profitability:** Ensure the price is able to cover the costs associated with acquiring, developing, marketing as to generate the desired profit margin.
        - **Market Acceptance:** Price should match what the target audience expects without compromising on perceived value.
        - **Competitor Pricing:** Understand how competitors have priced similar games and their strategy behind their success.
        - **Sales Volume Projections:** Estimate the demand for the product at different price points.

- **Value Perception:** Estimate the perceived value of the game at different price points.

2. **a. What is the market size for the types of games Athena sells in 2019? How did you determine this number and which resources did you use?**

My estimate for the market size for the types of games Athena sells games in the premium PC market is approximately **$3.9 billion**. As stated in the Nielsen report, the global premium PC game market was around $5.2 billion for 2019. The case stated that Athena sells all of its games exclusively through Steam and Steam has a global market share of 75%. Combining these two figures (5.2 billion*0.75) we get 3.9 billion.



Background

Athena Softworks, Inc. is a video game developer and publisher. They specialize in premium role-playing games (RPGs) for PC play and have had published eight games to date. Athena distributes its games exclusively using Steam, the world's largest distributor of PC games, which takes up 75% of the global market share. It is early 2020 and Athena is considering acquiring a new game title and there are three candidate games.

b. What do you project the market size to be in 2020, ignoring COVID-19? Why?

I would expect the market size to increase to **$3.975 billion** because the report stated that more major titles will be released in 2020 and it also projected the revenues for the premium PC market to increase to $5.3 billion (see image above). Additionally, the report stated that the decrease in revenue for the premium PC market was because there were major releases in 2018 and none in 2019.

c. How would you expect COVID-19 to impact this market? (No analysis, just a thoughtful response.) Moving forward, you may ignore any effects of COVID-19 on the market

COVID-19 would have a positive impact on the premium game market because more people would be stuck indoors. People will have a lot more time to play games and use gaming as a means of destressing or escaping boredom. Therefore, it wouldn't be surprising to see a huge jump in revenue for the whole digital gaming sector.

Furthermore, since premium PC games are sold through Steam or other online distributors, the supply chain won't be heavily impacted.

3. **Your coworkers conducted a survey conducted with a sample of prior customers. The survey contained a series of 40 statements and participants rated these on a 7-point Likert scale; the survey statements may be found on page 4. As you investigate this data, may use any of the provided resources as a guide on substantive content (e.g., if you are not familiar with the video game industry), but note that the results may not map cleanly to these secondary resources.**

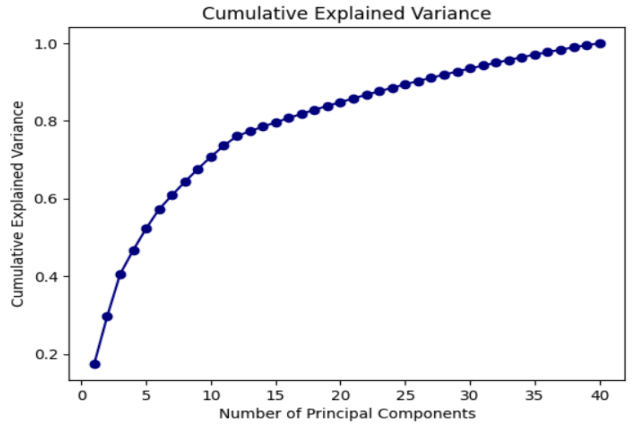
a. To identify potential segments in the market, first perform factor analysis. Include the factor loadings in your report. Name and interpret the factors. Identify the most relevant survey statements for each factor.

To determine if the factor analysis would be suitable, I ran Bartlett's test of Sphericity and a KMO test which both indicated that factor analysis would be suitable for this dataset. The p-value was <0.05 for Bartlett's test which indicated that the variables had some sort of correlation and would be suitable for factor analysis. Furthermore, the KMO test had a value >0.6 which indicated that 87% of the variables had a common variance.

```
# Step 1: Evaluate the data
# Bartlett's Test of Sphericity (we want p < 0.05 to go ahead)
from factor_analyzer.factor_analyzer import calculate_bartlett_sphericity
chi_square_value, p_value = calculate_bartlett_sphericity(factor_vars)
print(f"Bartlett's Test of Sphericity p-value: {p_value}")

# KMO-test (we want overall MSA > 0.6 to go ahead)
from factor_analyzer.factor_analyzer import calculate_kmo
kmo_all, kmo_model = calculate_kmo(factor_vars)
print(f"KMO-test overall MSA: {kmo_model}")
```

```
Bartlett's Test of Sphericity p-value: 0.0
KMO-test overall MSA: 0.8701380128313353
```




```
[6.99670203 4.88863646 4.35620131 2.45163899 2.23872323 1.99262917
1.44379198 1.39758246 1.31343252 1.25778586 1.13573613 0.99262882
0.49515349 0.47821435 0.44820805 0.44104524 0.43468685 0.41502891
0.40116737 0.3909749  0.38170652 0.37762322 0.36281221 0.36170064
0.35245708 0.34510457 0.33208262 0.32153735 0.31991887 0.31195317
0.30191319 0.29213323 0.28811272 0.27834207 0.27342393 0.26187468
0.24805693 0.23465604 0.21797911 0.21506989]
```

In order to determine the number of factors that explained the correlation between variables, I plotted a cumulative explained variance graph and looked at factors with eigenvalues larger than 1. The values showed that 11 factors were enough to explain 73% of the variance.

index	Fantasy and Immersion	Non-competitive Gameplay	Strategic Play	Action and Excitement	Non-Exploratory Play	Slow n Steady Play	Completion and Achievement	Anti-Narrative Engagement	Personalization	Cooperative Social Play	Anti-Challenge
imp.challenge	-0.0570654674	-0.03396996426	0.2776541858	-0.1263945895	-0.1649615798	0.02938757297	0.01075584055	0.09477014433	-0.02908818763	-0.02771075386	-0.731341006
imp.unlocks	0.0317739976	0.1307309983	-0.08805059215	-0.1147441128	0.01191884607	0.08998821134	0.8019336362	0.0576038948	0.0007757881873	-0.1799998473	-0.01061331654
imp.customize	0.1985793365	-0.06456331955	0.03972757933	-0.1351500456	-0.1151722092	-0.016080812	-0.006651127764	-0.1072462207	0.7724604993	0.009363363325	0.008098204724
imp.difficulty	-0.0176415704	-0.02676345911	0.2529030269	-0.1418707753	-0.1479624902	0.04579544233	0.01985444612	0.0867362726	0.02569929941	-0.07632499596	-0.6985295029
imp.characters	0.2681312036	0.09354421351	0.0273948497	0.0007195292943	-0.08775358683	0.02350351653	-0.03587730005	-0.7982388284	0.06089791084	-0.00750177979	0.05114584121
imp.storyline	0.2874494372	0.06196428839	0.03409617276	-0.01818160872	-0.09267315077	-0.009830085463	-0.06990508842	-0.7762103679	0.1317535964	-0.03734574287	0.07088933696
imp.mastery	-0.02752696802	-0.00372465654	0.5710131991	-0.1114186437	-0.1097711058	0.04171586527	-0.08894650494	0.004514962249	0.006972345537	0.03523241955	-0.4600592276
imp.backstory	0.2712394855	0.05174588104	0.000184410117	-0.004458994768	-0.09872884898	0.04902547884	-0.02290130955	-0.8023618139	0.1040608012	-0.03285573051	0.06744818785
imp.dominate	-0.01815347915	-0.7288112024	0.03093780774	0.002702553764	-0.05237528797	-0.1575297034	-0.1073011812	0.05058979738	0.06983729342	0.1214810554	-0.02573017815
imp.completion	0.01529818997	0.1611845741	-0.0552746981	-0.1233501526	0.04043038557	0.1064801394	0.7603716672	0.04734332833	-0.02634675454	-0.2016105383	0.01297897176
imp.wealth	0.1816936104	0.04887499628	0.03713665132	-0.06049734777	-0.1329803959	-0.1217872424	0.1056499135	-0.04466250573	0.2064556342	-0.3069096377	-0.0243636488
imp.fantasy	0.7643848561	0.04425578129	0.05340008966	-0.1158988299	-0.1254384747	0.005148873288	-0.011048565	-0.1218469199	0.1414538211	-0.035296651	0.01073548342
imp.items	0.6830932235	0.0479300956	0.03408058956	-0.07644021323	-0.1130167096	0.007695273289	9.68E-06	-0.3293279326	0.1192857861	-0.07512436541	-0.008222806892
imp.power	0.7471216337	0.01639607075	-0.01994325389	-0.106208877	-0.1106809695	0.05045582513	0.03122277528	-0.1433015924	0.1707191604	-0.03561370612	0.005346013321
imp.offbeat	0.1786103511	-0.004553639732	0.08633707111	-0.08856851905	-0.7343276527	0.01169902742	-0.04418253696	-0.1069420489	0.03774195592	0.0286041773	-0.1474940726
imp.collect	0.03122616289	0.1573527695	-0.06850566048	-0.116729683	0.03862725994	0.1248169963	0.7927843464	0.01949265305	-0.0001488463935	-0.1900062462	-0.003461476371
enj.excitement	0.001382957736	-0.1168653244	0.03328523896	0.1342469576	-0.04097349744	-0.7291881775	-0.06842109508	0.04412973071	0.05218224266	0.05137673723	0.02470822472
enj.destruction	-0.1325402251	-0.0278923826	-0.1046787494	0.7810057552	0.1151269137	-0.08757737485	-0.09165983406	-0.008452094021	-0.0822114892	0.01901138748	0.08409973899
enj.others	-0.004120537457	-0.3207704602	0.004517979068	0.01759719666	-0.04807445112	-0.05535967481	-0.1803523969	0.01640015793	0.0731293428	0.714325158	0.03361079087
enj.react	0.0009029074031	-0.1196052504	0.02612041634	0.08908959497	0.01886050698	-0.7562624071	-0.1190042118	0.03088969904	0.006257694769	-0.01378639704	0.00435716522
enj.duels	-0.07929075598	-0.7360080587	-0.02018736198	0.07100385978	0.0002727673289	-0.1108833245	-0.1816745028	0.07817295018	0.02275250821	0.2144449055	0.002115605304
enj.strategy	0.00827272154	-0.04914594195	0.8184492782	-0.07494241287	-0.1065667615	-0.03338877198	-0.05269037669	-0.01908093027	0.01525564978	0.006190817501	-0.0823302727
enj.roleplay	0.7451753079	0.02505607816	-0.01428483365	-0.07267117844	-0.1530207004	-0.0245434761	0.04722574105	-0.1627037268	0.153254424	-0.01038958056	0.04888308759
enj.competition	-0.06223171632	-0.7770878188	-0.01820241833	0.06549806762	0.005935069064	-0.1175170636	-0.1504247625	0.07342750131	-0.002800353928	0.2636474085	-0.02312455987
enj.decisions	0.03602462568	0.01231243288	0.7589010968	-0.1275016085	-0.08233103795	0.01885265293	-0.03368647509	-0.04390610078	0.001375700283	0.01961636247	-0.1207190389
enj.common.goa	0.0007796121865	-0.1635901706	0.02263355899	0.02612458336	-0.1030250357	-0.0643806564	-0.1500650079	0.0104267608	0.05618539635	0.7948174017	0.03655457909
enj.planning	0.014088962	0.02220159186	0.7890958613	-0.118570907	-0.09602839143	-0.04182661542	-0.06614316592	-0.01424501734	0.06614342997	0.002341682505	-0.09956121109
enj.immersion	0.7840464652	0.0346162327	-0.01009107431	-0.09507093826	-0.1374526611	-0.02566609079	0.01824688919	-0.1420938574	0.1347994851	-0.0398540292	0.01976913863
enj.helping	-0.02323032405	-0.1136589319	0.0704667388	0.005828579527	-0.06860867724	-0.06341641571	-0.143070596	0.01158062241	0.01441277546	0.7856660776	-0.001803002852
enj.fast	-0.005304558247	-0.0966896046	-0.03842436908	0.110193106	-0.02277078791	-0.7063282829	-0.0803412325	-0.01944609355	0.04794555239	0.0221544012	0.04534526202
enj.guns	-0.08980038844	-0.03227694694	-0.1297978639	0.7718194233	0.1008361172	-0.0919906418	-0.0963793099	0.008143376765	-0.07984791835	0.05366693686	0.07956993397
enj.gore	-0.09177779772	-0.03654555871	-0.1140900165	0.7809009845	0.111872863	-0.1154611225	-0.1367686554	0.02321336558	-0.1015120229	0.02587197691	0.09462519763
enj.blow.up	-0.1256945543	-0.04404311929	-0.126382078	0.7602317435	0.08000342673	-0.1052744822	-0.03759679031	0.01149599452	-0.133356234	0.03918847743	0.07622946219
freq.explore	0.1506004959	-0.009936825515	0.130526216	-0.08186154736	-0.7663576687	-0.01245979369	-0.05694121007	-0.0633997639	0.111780885	0.02789388491	-0.05929763131
freq.experiment	0.1153415691	0.002385622874	0.1195322508	-0.1208924526	-0.7833720808	-0.01298527377	-0.02856387632	-0.04126470925	0.06058474159	0.06703921873	-0.05645020722
freq.study	0.003264514705	0.03444932637	0.5640660458	-0.1148735648	-0.1031107409	0.009313590083	0.003837477321	0.02519890477	0.006002609096	-0.02419276188	-0.4855922454
freq.char.creatio	0.204525734	-0.03689606813	0.03051241483	-0.09381706605	-0.08742619707	-0.01789518328	-0.01792341047	-0.0827292154	0.7325409499	0.009577527349	-0.008233412184
freq.stats	0.183210835	0.05743572794	0.1118349937	-0.1095980278	-0.09458430259	-0.1507516185	0.1012216361	-0.03668824014	0.1591045543	-0.2863758053	-0.04342636915
freq.customize	0.2188279324	-0.0133275884	0.002408605658	-0.1273152418	-0.07084246662	-0.07699911513	-0.01906728638	-0.07714651841	0.739975535	-0.0004650532128	0.003511074813
freq.test.world	0.1528528442	-0.04999309359	0.08338570097	-0.1054553429	-0.7870767707	-0.04837297151	0.04025328772	-0.06795398767	0.1038684618	-0.01347319921	-0.1022476847

Significant factor loadings are highlighted in green

There were a total of 11 factors identified:

Factor 1 - Fantasy and Immersion: This factor has high loadings in imp.fantasy, imp.items, imp.power, enj.roleplay and enj.immersion. Gamers with high loadings like fantasy, obtaining powerful items and immersing themselves in a different world. They also like customization, exploring the world and storylines. These gamers would love RPG games.

Factor 2 - Non-competitive Gameplay: This factor has high negative loadings for imp.dominate, enj.duels and enj.competition. Other loadings also indicate that this type of gamer prefers solo play (-0.32 for enj.others). High scores in this factor would mean these gamers prefer a more relaxing completion based game rather than competing against other players. These gamers would probably prefer Arcade-like games (Candy-Crush, Animal Crossing)

Factor 3 - Strategic Play: This factor has high loadings in imp.strategy, enj.strategy, enj.decisions, enj.planning and freq.study. High scores in this factor point to a gamer who likes strategic thinking, planning and study of other players. Gamers with high scores in this factor are analytical, often studying other players to improve their own gameplay. They enjoy games that require strategic thinking and skill development. Strategy games like Civilisation or even Chess would suit this type of gamer.

Factor 4 - Action and Excitement: This factor has high loadings in enj.destruction, enj.guns, enj.gore and enj.blow.up. As the name suggests, high scores in this factor mean the gamer enjoys adrenaline-packed action games. First-Person Shooters like Call-of-Duty, PUBG or Valorant would be the perfect games.

Factor 5 - Non-exploratory Play: Dominated by negative loadings in imp.offbeat, freq.explore, and freq.experiment. This factor represents gamers who dislike exploring new, unconventional ways of playing and discovering various aspects of the game world. They are likely to be attracted to games that are well-structured and conventional in gameplay.

Factor 6 - Slow n Steady: This factor has low scores in enj.excitement, enj.react and enj.fast potentially meaning a preference for slower-paced games. Puzzle games and games which require a lot of thinking could be good fits for these types of gamers.

Factor 7 - Completion and Achievement: This factor is characterized by imp.unlocks and imp.completion. Gamers scoring high in this factor are probably driven by the desire

to complete every aspect of a game and unlock all achievements. They find satisfaction in completing tasks and collecting items.

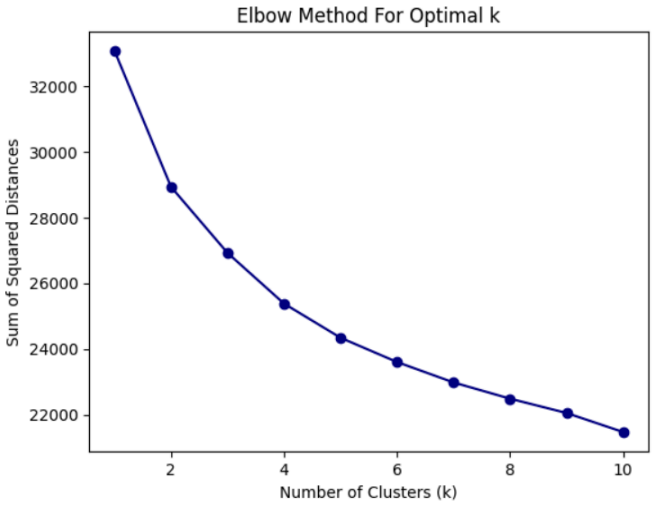
Factor 8 - Anti-narrative Engagement: Imp.characters, imp.storyline, and imp.backstory are prominent here. This factor is about the importance of story and character development in games. Gamers with high scores dislike games with rich narratives and complex characters.

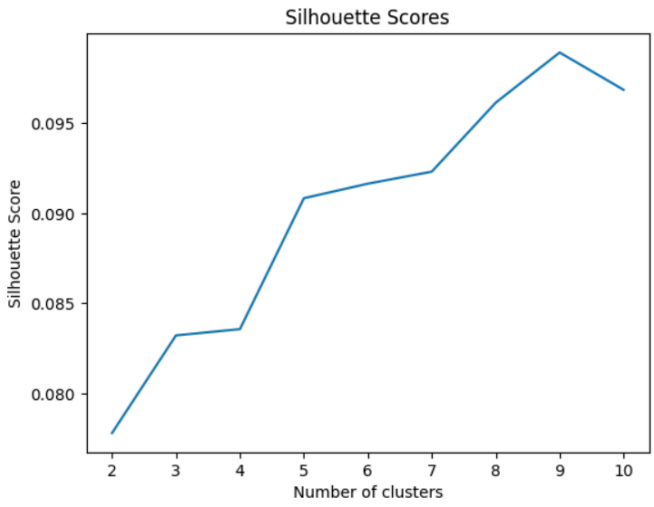
Factor 9 - Personalization and Customization: High loadings in imp.customize, freq.char.creation, and freq.customize indicate a preference for personalizing gaming experiences. These gamers value the ability to customize characters and settings, enjoying games that offer detailed customization options.

Factor 10 - Cooperative Social Play: Marked by enj.others, enj.common.goal, and enj.helping. Gamers who score high on this factor enjoy playing with others, cooperating towards common goals, and helping fellow players. Open world interactive games like Massively Multiplayer Online (MMO) games would suit this type of gamer.

Factor 11 - Anti-Challenge and Anti-Mastery: Imp.challenge and imp.difficulty have significant negative scores. These indicate a preference for easy, straight-forward gameplay and mastery. These gamers don't like overcoming difficult challenges and playing games at higher difficulty levels.

b. Next, perform cluster analysis using K-means clustering to identify segments. Include the cluster centers in your report. Name and interpret these segments based on these cluster centers. Identify the most relevant factors for each segment.





To determine the number of clusters suitable for segmenting gamers, I used the elbow method and average silhouette scores. As the elbow method didn't show a clear optimal amount of clusters I relied on the plot of the average silhouette scores. Although the plot of the silhouette scores showed that 9 clusters would have the maximum representation of the variables, I found that 5 segments would have the best results and showed maximum separation between the clusters. This was also because the increase in separation between clusters was not as big between 1-5 vs. 5-9.

Additionally, after segmenting the gamers into 8 clusters, there was a certain amount of overlap in the Fantasy & Immersion factor. The only difference between these two clusters was the preference for personalization vs narrative.

index	Fantasy&Immersion	Non-competitive	Strategic Play	Action&Excitement	Non-Exploratory	Slow&Steady Play	Achievement&Completion	Anti-Narrative	Customization&Creation	Cooperative Social Play	Anti-Challenge
Cluster 0	1.019011273	-0.4161658478	-0.04300218265	-0.02324884528	-0.4926580171	-0.4252055059	-0.05568721449	-0.1957361328	0.08757960276	-0.1034109057	-0.03612882062
Cluster 1	-0.4239576215	-0.2689723735	1.19728568	0.08631269742	0.1400346913	0.0122438775	-0.3468990744	0.2499797235	-0.1251653955	0.1543337321	0.000232250017
Cluster 2	-0.8746925008	-0.004126231087	-0.7270902389	0.9444571289	-0.0974460306	-0.2477783287	0.1760161711	-0.0777383184	-0.1495031063	-0.0155161691	-0.03836939326
Cluster 3	-0.1321879116	-0.6242272323	-0.8254381743	-0.612940308	0.3106333319	0.8047598829	-0.04718430394	0.1741308734	0.3636410439	-0.02031163791	0.02732878291
Cluster 4	0.170979073	1.216265191	0.08487825894	-0.4049293032	0.2423192202	0.0480676232	0.3006009441	-0.1079516017	-0.1192568831	-0.01534906029	0.0514518179

Cluster 0 - "Immersive Adventurers": High in Fantasy&Immersion but low in Non-competitive Gameplay and Action&Excitement. These gamers likely prefer rich, immersive worlds and story-driven experiences, enjoying games more for their narrative and environment than for competition or fast-paced action. Role-Playing Games would be a good fit for these types of gamers.

Cluster 1 - "Strategic Competitors": Characterized by high Strategic Play and moderate Action&Excitement. These gamers enjoy games that require strategic thinking and planning. They might be drawn to competitive games where strategic skills are crucial, but they also appreciate a moderate level of action and dynamism in gameplay.

Cluster 2 - "Dynamic Action Seekers": With high scores in Action&Excitement and low in Strategic Play and Slow&Steady Play, this group likely prefers fast-paced, action-oriented games. They enjoy dynamic and thrilling gameplay, with less emphasis on strategic planning or immersive narratives. FPS games like Valorant, CounterStrike GO and PUBG would be suitable games for these gamers.

Cluster 3 - "Relaxed Enthusiasts": Marked by low Action&Excitement and high in Slow&Steady Play and Non-Exploratory Play. These gamers might prefer more relaxed, well-structured games like puzzle games where they can take their time and enjoy the experience at their own pace. Casual games like Candy Crush or Animal Crossing would be good games for this segment.

Cluster 4 - "Community Engagers": High in Non-competitive Gameplay and Cooperative Social Play. These gamers are likely drawn to games that offer rich social interactions and cooperative gameplay experiences. They might prefer playing in a community-focused environment, valuing collaboration and social aspects of gaming over competition.

c. Finally, use cross tabulation and regression analysis to investigate the relationships between the segments and various demographic attributes (gender, age, income, location). Identify any significant relationships and describe each of the resulting segments in terms of their demographic attributes (% female, average age, and average income) regardless of statistical significance.

In order to incorporate demographics into the analysis, I used cross tabulation and ran chi-squared tests of independence on the demographic attributes and the segments to see if there were any significant observations. I also ran regression analysis on the age and income attributes since these were nominal variables.

Gender and Segments:

Chi-squared value: 12.03594195644825
P-value: 0.1496074314820157

gender	Segment_Label	Community Engagers	Dynamic Action Seekers	Immersive Adventurers	Relaxed Enthusiasts	Strategic Competitors
female	Observed	73	65	97	71	71
	Expected	79.32	68.38	87.07	61.09	81.14
	Chi squared	0.5	0.17	1.13	1.61	1.27
male	Observed	99	82	91	63	102
	Expected	91.94	79.26	100.93	70.81	94.06
	Chi squared	0.54	0.09	0.98	0.86	0.67
nonbinary	Observed	2	3	3	0	5
	Expected	2.74	2.36	3.0	2.11	2.8
	Chi squared	0.2	0.17	0.0	2.11	1.73

Based on this crosstab, there doesn't seem to be a strong enough statistical relationship between gender and the segments. The overall p-value was 0.1496 which is higher than the typical 0.05. Furthermore, the critical value for the chi-squared statistic with 8 degrees of freedom at the 0.05 significance level is approximately 15.51. The value for this crosstab was only 12.04 which falls below the critical value and indicates that there is no significant relationship between Gender and the segments.

The largest chi-squared contributions come from the "Relaxed Enthusiasts" segment for nonbinary (2.11) and "Strategic Competitors" segment for nonbinary (1.73). This might indicate that these cells are the furthest from what would be expected if there were no association, but since the overall p-value is not below 0.05, these deviations alone are not enough to conclude a significant association overall.

Age and Segments:

Chi-squared value: 106.75709784006334
P-value: 7.635555188373134e-14

ageNew	Segment_Label	Community Engagers	Dynamic Action Seekers	Immersive Adventurers	Relaxed Enthusiasts	Strategic Competitors
18–25	Observed	48	79	95	82	104
	Expected	85.84	74.0	94.23	66.11	87.82
	Chi squared	16.68	0.34	0.01	3.82	2.98
26–35	Observed	68	43	63	48	61
	Expected	59.54	51.33	65.36	45.85	60.91
	Chi squared	1.2	1.35	0.09	0.1	0.0
36–45	Observed	23	21	20	4	6
	Expected	15.57	13.42	17.09	11.99	15.93
	Chi squared	3.55	4.28	0.5	5.32	6.19
46-55	Observed	12	4	8	0	5
	Expected	6.1	5.26	6.7	4.7	6.24
	Chi squared	5.7	0.3	0.25	4.7	0.25
56-65	Observed	22	3	5	0	2
	Expected	6.73	5.8	7.39	5.19	6.89
	Chi squared	34.62	1.35	0.77	5.19	3.47
65 and above	Observed	1	0	0	0	0
	Expected	0.21	0.18	0.23	0.16	0.22
	Chi squared	2.96	0.18	0.23	0.16	0.22

Based on the crosstab of age and segments, there seems to be a strong statistically significant relationship because the p-value is very small and well under 0.05. The Chi-squared value of **106.76** is also much higher than the critical value of **31.41** for this cross tabulation.

Additionally, the observed counts in some cells of the table are notably different from the expected counts, contributing to high chi-squared values for certain age groups and segments. There are a lot more 18-25 year olds who fall under the ‘Community Engagers’ segment contributing to a chi-squared value of **16.68** for that cell. Similarly, the 56-65 age group also shows a significant deviation from the expected count in the 'Community Engagers' segment, with a chi-squared value of **34.62** for that cell. These large discrepancies between observed and expected counts across several cells support the conclusion of a significant association between age group and segment label.

Regression Analysis: Age and Segments

OLS Regression Results

Dep. Variable:	age	R-squared:	0.124
Model:	OLS	Adj. R-squared:	0.119
Method:	Least Squares	F-statistic:	28.97
Date:	Sat, 16 Dec 2023	Prob (F-statistic):	1.50e-22
Time:	18:09:47	Log-Likelihood:	-2993.1
No. Observations:	827	AIC:	5996.
Df Residuals:	822	BIC:	6020.
Df Model:	4		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	33.6897	0.686	49.077	0.000	32.342	35.037
Dynamic Action Seekers	-6.8030	1.009	-6.743	0.000	-8.783	-4.823
Immersive Adventurers	-6.7054	0.949	-7.066	0.000	-8.568	-4.843
Relaxed Enthusiasts	-9.9285	1.041	-9.540	0.000	-11.971	-7.886
Strategic Competitors	-8.3807	0.965	-8.682	0.000	-10.275	-6.486

Omnibus:	227.461	Durbin-Watson:	1.878
Prob(Omnibus):	0.000	Jarque-Bera (JB):	524.900
Skew:	1.478	Prob(JB):	1.05e-114
Kurtosis:	5.548	Cond. No.	5.72

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

The regression analysis shows that age significantly differs across customer segments, with each segment typically being younger than the ‘Community Engagers’ segment. ‘Relaxed Enthusiasts’ are notably the youngest. All segment coefficients are statistically significant, as indicated by the p-values being less than 0.05. The model explains approximately 11.9% of the variance in age, which means the relationship is probably not linear.

State and Segments:

Chi-squared value: 11.584092636923902
P-value: 0.479632508354895

stateNew	Segment_Label	Community Engagers	Dynamic Action Seekers	Immersive Adventurers	Relaxed Enthusiasts	Strategic Competitors
Midwest	Observed	35	24	37	24	35
	Expected	32.61	28.11	35.8	25.11	33.36
	Chi squared	0.17	0.6	0.04	0.05	0.08
Northeast	Observed	27	25	31	30	35
	Expected	31.14	26.84	34.18	23.98	31.85
	Chi squared	0.55	0.13	0.3	1.51	0.31
South	Observed	75	66	65	45	68
	Expected	67.12	57.86	73.67	51.69	68.66
	Chi squared	0.93	1.15	1.02	0.87	0.01
West	Observed	37	35	58	35	40
	Expected	43.13	37.18	47.35	33.22	44.12
	Chi squared	0.87	0.13	2.4	0.1	0.39

After running a crosstab assessing the association between U.S. regions (Midwest, Northeast, South, West) and customer segments, the crosstab suggested no statistically significant difference between the expected and observed gaming segments and regions. The chi-squared statistic is 11.58 with a p-value of 0.48, which exceeds the typical threshold of 0.05. Essentially, regional location does not appear to influence how gamers are distributed among the segments. This is consistent across all regions, as indicated by the low chi-squared contributions for each cell.

Income and Segments:

Chi-squared value: 61.38537727487949
P-value: 0.00027016371696675346

incomeNew	Segment_Label	Community Engagers	Dynamic Action Seekers	Immersive Adventurers	Relaxed Enthusiasts	Strategic Competitors
Under $20,000	Observed	10	24	35	26	30
	Expected	26.3	22.67	28.87	20.25	26.9
	Chi squared	10.1	0.08	1.3	1.63	0.36
$20,000 to $39,999	Observed	41	48	47	47	57
	Expected	50.5	43.53	55.43	38.89	51.66
	Chi squared	1.79	0.46	1.28	1.69	0.55
$40,000 to $59,999	Observed	35	28	42	33	46
	Expected	38.71	33.37	42.5	29.81	39.6
	Chi squared	0.36	0.87	0.01	0.34	1.03
$60,000 to $79,999	Observed	37	21	27	18	20
	Expected	25.88	22.31	28.41	19.93	26.47
	Chi squared	4.78	0.08	0.07	0.19	1.58
$80,000 to $99,999	Observed	25	16	16	7	10
	Expected	15.57	13.42	17.09	11.99	15.93
	Chi squared	5.71	0.5	0.07	2.08	2.21
$100,000 to $119,999	Observed	11	7	14	1	10
	Expected	9.05	7.8	9.93	6.97	9.26
	Chi squared	0.42	0.08	1.67	5.11	0.06
$120,000 to $149,999	Observed	9	5	8	1	5
	Expected	5.89	5.08	6.47	4.54	6.03
	Chi squared	1.64	0.0	0.36	2.76	0.17
$150,000 and above	Observed	6	1	2	1	0
	Expected	2.1	1.81	2.31	1.62	2.15
	Chi squared	7.21	0.37	0.04	0.24	2.15

The chi-squared test results show a value of 61.39 with a p-value of 0.0027, indicating a statistically significant association between income levels and gamer segments. Specifically, the number of gamers in certain income brackets deviates more than expected from the numbers that would be seen if there were no relationship between income and segment. For example, there are fewer gamers than expected earning under $20,000 in most segments, while in the $150,000 and above bracket, there are more gamers than expected across the segments. This suggests income level is likely a factor in segment distribution.

Regression Analysis: Income and Segments

OLS Regression Results

Dep. Variable:	income	R-squared:	0.053
Model:	OLS	Adj. R-squared:	0.049
Method:	Least Squares	F-statistic:	11.56
Date:	Sat, 16 Dec 2023	Prob (F-statistic):	3.92e-09
Time:	18:09:47	Log-Likelihood:	-9773.1
No. Observations:	827	AIC:	1.956e+04
Df Residuals:	822	BIC:	1.958e+04
Df Model:	4		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	6.485e+04	2495.093	25.991	0.000	6e+04	6.97e+04
Dynamic Action Seekers	-1.382e+04	3667.022	-3.770	0.000	-2.1e+04	-6626.078
Immersive Adventurers	-1.174e+04	3449.185	-3.404	0.001	-1.85e+04	-4970.381
Relaxed Enthusiasts	-2.307e+04	3782.767	-6.100	0.000	-3.05e+04	-1.56e+04
Strategic Competitors	-1.912e+04	3508.714	-5.449	0.000	-2.6e+04	-1.22e+04

Omnibus:	186.661	Durbin-Watson:	1.960
Prob(Omnibus):	0.000	Jarque-Bera (JB):	422.370
Skew:	1.217	Prob(JB):	1.92e-92
Kurtosis:	5.517	Cond. No.	5.72

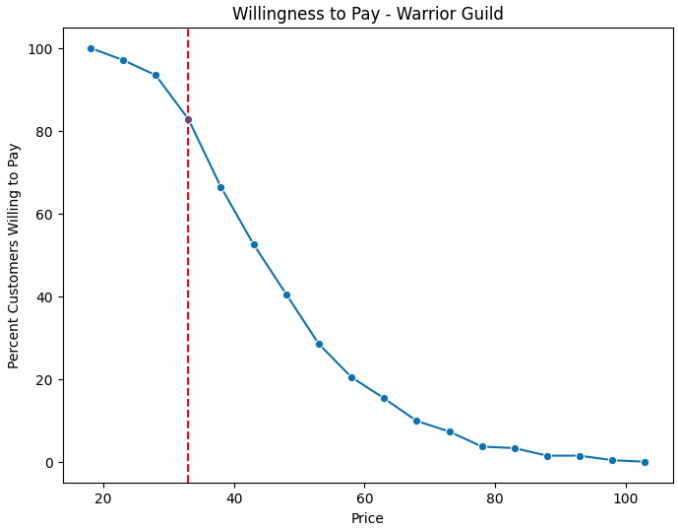
Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

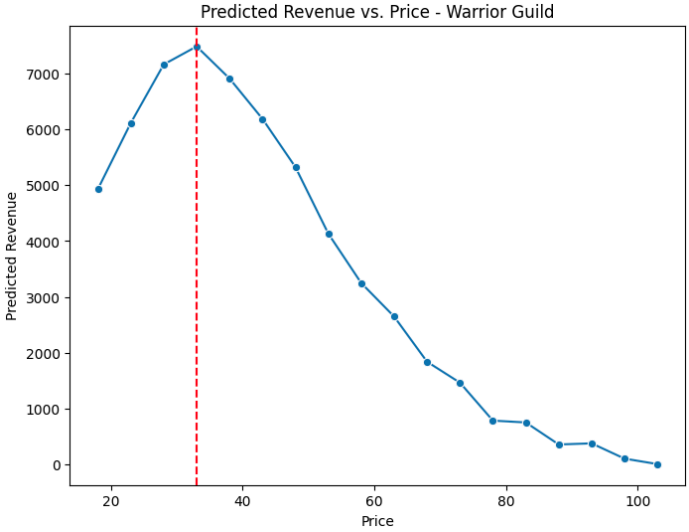
The regression analysis demonstrates significant disparities in income across gaming segments relative to 'Community Engagers'. Segments 'Dynamic Action Seekers', 'Immersive Adventurers', 'Relaxed Enthusiasts', and 'Strategic Competitors' report lower income levels, as evidenced by the negative coefficients. Despite the statistical significance of the model, its explanatory power is very limited (R-Squared: 0.05), with segment variables accounting for a small portion of income variation.

4. [15pt] Next, investigate another part of the survey: Gabor Granger responses for each game. Each respondent was randomly presented with one of the three games and the survey identified the maximum price point at which each respondent would "probably purchase" the presented game.

a. For each game, show the two Gabor Granger plots: percent customers willing to pay and predicted revenue as a function of price. What is the ideal price point for each game?

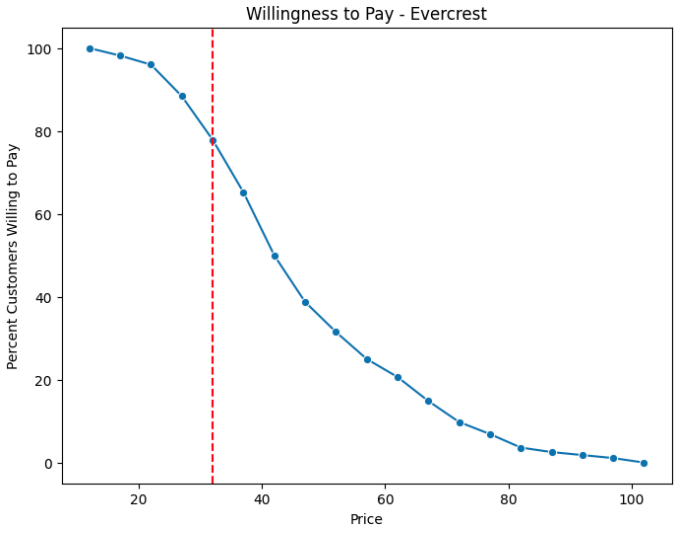
Warrior Guild:

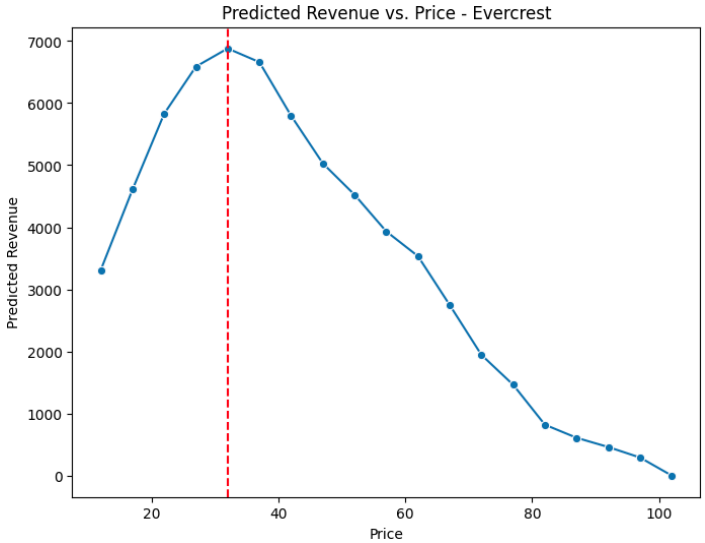




The ideal price point for Warrior Guild is $33. At this price point, 82.85% of gamers are willing to pay for Warrior Guild and the predicted maximum revenue is $7491.

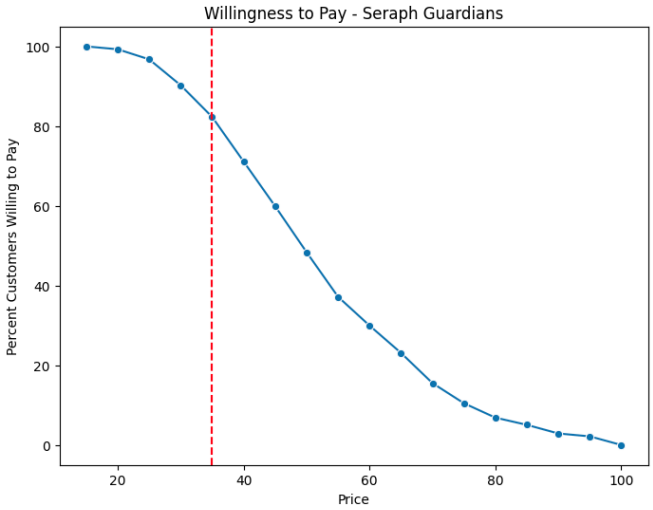
Evercrest:





The ideal price for Evercrest is $32 and about 77.90% of gamers are willing to pay this price. The predicted maximum revenue at this price will be $6880.

Seraph Guardians:



The ideal price for Seraph Guardians is $35 and about 82.31% of gamers are willing to pay this price. The predicted maximum revenue at this price will be $7980.

b. Use linear regression to predict which segment is most interested in each game (willing to pay the most). Which segments are most and least interested in each game?

The regression analysis below predicts how much each segment is willing to pay for a game based on the coefficient of each segment.

Warrior Guild:

OLS Regression for Warrior Guild:

OLS Regression Results

Dep. Variable:	gg.maxprice	R-squared:	0.023
Model:	OLS	Adj. R-squared:	0.008
Method:	Least Squares	F-statistic:	1.569
Date:	Sat, 16 Dec 2023	Prob (F-statistic):	0.183
Time:	22:43:37	Log-Likelihood:	-1136.2
No. Observations:	274	AIC:	2282.
Df Residuals:	269	BIC:	2301.
Df Model:	4		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	38.3609	0.784	48.923	0.000	36.817	39.905
Community Engagers	8.5855	1.858	4.620	0.000	4.927	12.244
Dynamic Action Seekers	9.8782	2.018	4.896	0.000	5.906	13.850
Immersive Adventurers	9.6391	1.739	5.544	0.000	6.216	13.062
Relaxed Enthusiasts	3.2224	1.982	1.626	0.105	-0.679	7.124
Strategic Competitors	7.0356	1.832	3.841	0.000	3.429	10.642

Omnibus:	31.102	Durbin-Watson:	1.904
Prob(Omnibus):	0.000	Jarque-Bera (JB):	38.165
Skew:	0.857	Prob(JB):	5.16e-09
Kurtosis:	3.635	Cond. No.	3.71e+15

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
[2] The smallest eigenvalue is 2.39e-29. This might indicate that there are
strong multicollinearity problems or that the design matrix is singular.

Based on the regression for Warrior Guild, the most interested segment (segment willing to pay the most) would be the ‘Dynamic Action Seekers’ and the least interested segment would be the ‘Relaxed Enthusiasts’.

Evercrest:

OLS Regression for Evercrest:

OLS Regression Results

Dep. Variable:	gg.maxprice	R-squared:	0.037
Model:	OLS	Adj. R-squared:	0.023
Method:	Least Squares	F-statistic:	2.594
Date:	Sat, 16 Dec 2023	Prob (F-statistic):	0.0369
Time:	22:43:37	Log-Likelihood:	-1182.4
No. Observations:	276	AIC:	2375.
Df Residuals:	271	BIC:	2393.
Df Model:	4		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	38.0683	0.898	42.405	0.000	36.301	39.836
Community Engagers	10.9317	2.180	5.014	0.000	6.639	15.224
Dynamic Action Seekers	6.7765	2.101	3.226	0.001	2.641	10.912
Immersive Adventurers	11.4650	2.072	5.534	0.000	7.386	15.544
Relaxed Enthusiasts	6.4555	2.406	2.684	0.008	1.719	11.192
Strategic Competitors	2.4396	2.031	1.201	0.231	-1.560	6.439

Omnibus:	24.031	Durbin-Watson:	2.193
Prob(Omnibus):	0.000	Jarque-Bera (JB):	28.045
Skew:	0.775	Prob(JB):	8.13e-07
Kurtosis:	3.187	Cond. No.	7.59e+15

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
[2] The smallest eigenvalue is 5.76e-30. This might indicate that there are
strong multicollinearity problems or that the design matrix is singular.

Based on the regression for Evercrest, the most interested segment (segment willing to pay the most) would be the ‘Immersive Adventurers’ and the least interested segment would be the ‘Strategic Competitors’.

Seraph Guardians:

OLS Regression for Seraph Guardians:

OLS Regression Results

Dep. Variable:	gg.maxprice	R-squared:	0.138
Model:	OLS	Adj. R-squared:	0.126
Method:	Least Squares	F-statistic:	10.90
Date:	Sat, 16 Dec 2023	Prob (F-statistic):	3.25e-08
Time:	22:43:37	Log-Likelihood:	-1166.6
No. Observations:	277	AIC:	2343.
Df Residuals:	272	BIC:	2361.
Df Model:	4		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	41.8477	0.836	50.044	0.000	40.201	43.494
Community Engagers	16.1369	1.866	8.648	0.000	12.463	19.811
Dynamic Action Seekers	5.8914	2.152	2.738	0.007	1.655	10.128
Immersive Adventurers	15.7215	1.866	8.425	0.000	12.048	19.395
Relaxed Enthusiasts	3.0159	2.193	1.375	0.170	-1.302	7.334
Strategic Competitors	1.0821	1.968	0.550	0.583	-2.792	4.956

Omnibus:	12.629	Durbin-Watson:	2.072
Prob(Omnibus):	0.002	Jarque-Bera (JB):	13.436
Skew:	0.539	Prob(JB):	0.00121
Kurtosis:	3.020	Cond. No.	5.16e+15

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
[2] The smallest eigenvalue is 1.25e-29. This might indicate that there are
strong multicollinearity problems or that the design matrix is singular.

The regression for Seraph Guardians show that the most interested segment (segment willing to pay the most) are the 'Community Engagers' tightly followed by 'Immersive Adventurers' and the least interested segment would be the 'Strategic Competitors'.

c. Assume that only 30% of respondents who indicated they would "probably purchase" at a given price will actually do so within the first year. Also assume that the survey sample was representative of the approximately 10 million active Steam customers who have expressed interests in similar types of games. What would be the gross and net revenues for each game in the first year?

	No. of Players	Percentage of players	Real Percentage	Real Buyers	Ideal Price	Gross Revenue	
Warrior Guild	10000000	0.828467	30%	2485401	$ 33.00	$ 82,018,233.00	
Seraph Guardians	10000000	0.823105	30%	2469315	$ 35.00	$ 86,426,025.00	
Evercrest	10000000	0.778986	30%	2336958	$ 32.00	$ 74,782,656.00	
	Gross Revenue	Royalties Paid out	Cost of Game + Additional Costs	Valve Cost <10m	Valve Cost 10m-50m	Valve Cost >50m	Net Revenue
Warrior Guild	$ 82,018,233.00	$ (4,100,911.65)	$ (12,000,000.00)	$ (3,000,000.00)	$ (12,500,000.00)	$ (6,403,646.60)	$ 44,013,674.75
Seraph Guardians	$ 86,426,025.00	$ (4,321,301.25)	$ (12,500,000.00)	$ (3,000,000.00)	$ (12,500,000.00)	$ (7,285,205.00)	$ 46,819,518.75
Evercrest	$ 74,782,656.00	$ (3,739,132.80)	$ (13,000,000.00)	$ (3,000,000.00)	$ (12,500,000.00)	$ (4,956,531.20)	$ 37,586,992.00

The gross revenue for **Warrior Guild** is ~**$82 million** and the net revenue would be ~**$44 million**. The gross revenue for Seraph Guardians is ~**$86.4 million** and the net revenue would be ~**$46.82 million**. The gross revenue for Evercrest is ~**$74.78 million** and the net revenue would be ~**$37.59 million**.

5. [10pt] The final portion of the part of the survey asked respondents to rank six games with 1 being the most preferred choice. The six games include the three candidate games and three games that competitors have already announced will be on the market.

a. Assuming all the games are priced equally, that the surveyed customers are representative of the market, and that each customer purchases only one game, calculate the percentage of the market share Athena would have under each of the action alternatives.

```
{'rank.WarriorGuild': 11.970979443772672,
 'rank.Evercrest': 10.157194679564691,
 'rank.SeraphGuardians': 53.80894800483675}
```

The market share would be as follows:

Warrior Guild: 11.98%

Evercrest: 10.18%
Seraph Guardians: 53.81%

b. Discuss which of the assumptions above you might want to change, and in what ways, to generate more realistic estimates of market share under each of the action alternatives. Extra Credit: modify your simulation to actually change some or all of the assumptions you discuss and share the results.

Assumptions made above and what I would change:

Games are priced equally: Pricing strategies will vary and even with the Gabor Granger plots prices will change under different market conditions. The ideal prices calculated in the Gabor Granger plot should be used as starting prices and the customers willing to pay at the price should also be used as a proxy.

Surveyed customers are representative of the market: The survey may not be representative of the market. There may be biases present depending on how the survey was conducted and if random sampling or stratified random sampling was used. Using a larger, more diverse sample for the survey could help improve the representativeness. Alternatively, weighting responses to reflect the broader population could also help in cases where the sample is not perfectly representative.

Customers only purchase one game: Customers are more likely to purchase multiple games, especially if they fall under the same genre. A more realistic model would allow for the possibility of customers buying more than one game. This change would require considering the likelihood of multiple purchases and how preferences influence the combination of games purchased.

Exclusive preference for the highest ranked game: Gaming customers might consider the second or third ranked games if their most preferred games are not available or too expensive. Assigning probabilities to each rank could be a good way of calculating the market share.

Extra Credit:

I added weighted probabilities to the calculation and included the ideal price calculated from the Gabor Grangor plots. I also factored in the percent of willingness-to-pay customers for each game. Additionally, I changed the probability of buying a game if they ranked the game 1st from 1 to 0.3 to err on the conservative side.

The adjusted simulation, including competitor games and modified probabilities (where the top choice now has a probability of 0.3), yields the following estimated weighted market shares for all games:

```
# Adjusting the simulation to include competitor games with modified probabilities
rankings = data5.iloc[:,46:52]
# Adjusted probabilities for each rank
adjusted_rank_probabilities = {1: 0.3, 2: 0.21, 3: 0.15, 4: 0.09, 5: 0.06, 6: 0.03}

# Competitor games
competitor_games = ['DevilsGate', 'Marksman', 'QuestoftheTitan']
all_games = list(games_info.keys()) + competitor_games

# Initialize market share count for Athena's and competitor games
total_market_share = {game: 0 for game in all_games}

# Simulating weighted market share including competitor games
for idx, row in rankings.iterrows():
    for game in all_games:
        # Check if the game is Athena's game or a competitor's game
        if game in games_info:
            wtp_percentage = games_info[game]['wtp_percentage'] / 100
        else:
            wtp_percentage = 0.8  # Example default WTP for competitor games

        # Rank of the game for the respondent
        game_rank_key = f'rank.{game.replace(" ", "")}'
        game_rank = row[game_rank_key]

        # Probability of purchase based on the rank and willingness to pay
        purchase_probability = adjusted_rank_probabilities[game_rank] * wtp_percentage

        # Simulating the purchase decision
        if np.random.rand() < purchase_probability:
            total_market_share[game] += 1

# Calculating weighted market share percentage for all games
weighted_total_market_share_percentage = {game: (count / market_size) * 100 for game, count in total_market_share.items()}

weighted_total_market_share_percentage
```

```
{'Warrior Guild': 11.24546553808948,
 'Evercrest': 9.915356711003627,
 'Seraph Guardians': 20.314389359129382,
 'DevilsGate': 9.915356711003627,
 'Marksman': 6.045949214026603,
 'QuestoftheTitan': 11.124546553808948}
```

Warrior Guild: ~11.26% of the market share.
Evercrest: ~9.15% of the market share.
Seraph Guardians: ~20.31% of the market share.
DevilsGate (Competitor's Game): ~9.91% of the market share.
Marksman (Competitor's Game): ~6.04% of the market share.
QuestoftheTitan (Competitor's Game): ~11.12% of the market share.

This simulation suggests that Seraph Guardians would capture the largest market share among Athena's games, even with the inclusion of competitor games and the

adjusted probabilities for multiple game purchases. The competitor games also capture significant portions of the market, with DevilsGate showing a similar market share to Warrior Guild.

6. [10pt] Provide your final recommendations for each of the key decisions (part 1). As part of your recommendation on positioning, indicate whether you recommend targeting particular segment(s) or a non-targeting strategy. If you recommend a targeted approach, indicate which segment(s) should be targeted and justify your response. If you recommend a non-targeted approach, similarly justify your response.

Decision 1: Which Game Should Athena Pursue?

Recommendation: Acquire Seraph Guardians
Justification: Seraph Guardians has the highest predicted market share **(20.31%)** and shows the greatest interest from the segment "Community Engagers," closely followed by "Immersive Adventurers." Its projected gross and net revenues **($86.4m and $46.82m respectively)** are also the highest among the three options, making it the most viable option financially and market-wise.

Decision 2: How Should the Game Be Priced?

Recommendation: Implement Dynamic Pricing
Justification: Dynamic pricing allows Athena to adjust the price according to market demand, sales performance, and promotions. This strategy aligns with the varying willingness to pay across different segments and can help maximize revenue. I would recommend starting at the ideal price points determined by the Gabor Granger analysis **($35 for Seraph Guardians)** and adjusting based on market response is recommended.

Positioning Strategy:

Recommendation: Targeted Approach
Target Segments: Primarily "Community Engagers" and secondarily "Immersive Adventurers."
Justification: The regression analysis and cluster analysis reveal that these segments are most willing to pay for Seraph Guardians. These segments prefer cooperative social play and immersive, story-driven experiences, which should be the focus in marketing and game development.

Additional Considerations:
Monitor Competitor Activity: Stay aware of competitor strategies, especially those targeting similar segments.
Customer Demographics: Tailor marketing strategies to appeal to the primary demographic profile of the target segments, considering age, income, and gaming preferences.

By focusing on Seraph Guardians and adopting a targeted approach, Athena can effectively tap into the most promising market segments, align with strategic goals, and optimize revenue generation.