Gaming On-line

Il nostro lavoro si pone l'obiettivo di effettuare un’analisi relativa alle preferenze di gioco di possibili utenti per un servizio di gaming on-line così da poter creare una offerta che soddisfi la domanda di mercato italiana.

Fieldwork e Trattamento dei dati

- Target della ricerca sono i giocatori on-line
- Il questionario è composto da 3 tipologie di domande: domande sulle caratteristiche personali, domande sulle caratteristiche di gioco, domande sui servizi aggiuntivi utilizzati/utilizzabili dagli utenti
- Test sul questionario per verificarne la reale comprensione
- La raccolta dati è avvenuta tramite questionario on-line (creato tramite www.freeonlinesurvey.com), composto da domande quantitative e quantitative discrete
- Il questionario è stato sottoposto agli intervistati tramite link sui principali forum di gaming in Italia. (Multiplayer.it, TGMOnline.it, MMORPGItalia.it, Gamesradar.it, Gamesurf.it, IGZ.it, Ngi.it)
- Le 246 Survey individuali sono state successivamente ricopiate dal sito di supporto su un foglio excel

Analisi dei dati

L’analisi è composta da 7 parti principali:

- Analisi delle caratteristiche del target
- Analisi delle caratteristiche del core business (tipologia di gioco offerto, prezzo, sistema)
- Analisi dei servizi accessori offribili
- Modalità di pagamento
- Analisi fattoriale
- Regressione lineare
- Regressione logistica

Caratteristiche Target

Percezione sull'utilizzo di un nuovo servizio				
Creazione	Frequenza	Percentuale	Frequenza cumulata	Percentuale cumulata
No, utilizzo un servizio già esistente	41	16,67	41	16,67
Si, se gestito in modo efficiente	205	83,33	246	100

Sesso				
Sesso	Frequenza	Percentuale	Frequenza cumulata	Percentuale cumulata
Femmina	13	5,28	13	5,28
Maschio	233	94,72	246	100

Professione				
Professione	Frequenza	Percentuale	Frequenza cumulata	Percentuale cumulata
Studente	154	62,60	154	62,60
Impiegato	40	16,26	194	78,86
Libero Professionista	19	7,72	213	86,58
Studente - Lavoratore	15	6,10	228	92,68
Manager	8	3,25	236	95,93
Operaio	7	2,85	243	98,78
Disoccupato	3	1,22	246	100,00

E' inutile iniziare l'analisi se a priori non esiste una percezione da parte della possibile utenza sulla creazione e sull'utilizzo di un nuovo servizio di gaming italiano. L'83,33% degli intervistati ha dichiarato che utilizzerebbe un nuovo servizio se gestito in modo efficiente.

E' quindi possibile procedere con l'analisi delle caratteristiche del target di riferimento per creare un profilo di quelli che saranno i principali utenti del servizio.

L'utenza è quasi esclusivamente di sesso maschile, con un imponente 94,72% degli intervistati.

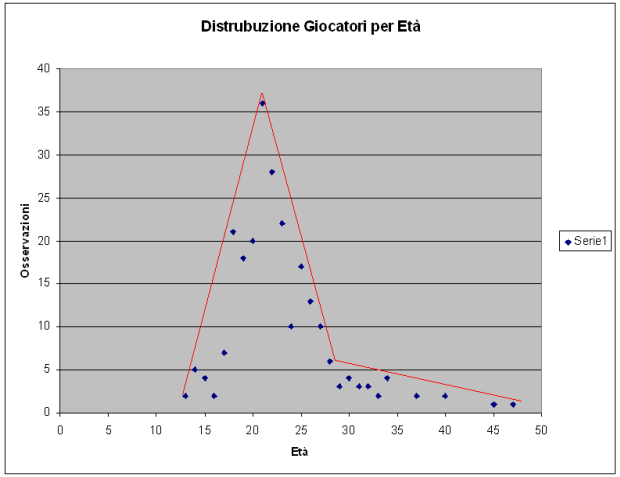
La professione invece è principalmente quella dello studente, con 62,60% rispetto al secondo, impiegato, con 16,26%. Per considerare una utenza più ampia bisogna considerare, però, la percentuale cumulata: se si vuole considerare solo l'80% della popolazione di riferimento, allora le professioni principali sono studente e impiegato; se si vuole arrivare fino al 90%, invece, bisogna aggiungere anche libero professionista e studente-lavoratore.

Età

Età			
Momenti			
N	246	**Somma dei pesi**	246,00
Media	22,72358	**Somma delle osservazioni**	5.590
Deviazione std	5,07704	**Varianza**	25,77634
Skewness	1,42942	**Kurtosis**	3,83733
SS non corretta	133.340	**SS corretta**	6.315,20325
Coeff variaz	22,34261	**Errore std media**	0,32370

Misure statistiche di base			
Posizione		**Variabilità**	
Media	22,72	**Deviazione std**	5,08
Mediana	22	**Varianza**	25,78
Moda	21	**Intervallo**	34
		Intervallo interquartile	5

Quantili	
Quantile	**Stima**
100% Max	47
99%	40
95%	32
90%	29
75% Q3	25
50% Mediana	22
25% Q1	20
10%	18
5%	16
1%	14
0% Min	13



L'età della popolazione di riferimento si assesta su una media di 22,72, con un valore centrale (mediana) di 22 e valore con più alta frequenza di 21 (moda). L'intervallo è di 24 anni (come indicano i quantili da 13 a 47) con un intervallo interquartile di 5 anni (da 20 a 25). Questi dati ci forniscono due indicazioni. La prima è la coerenza rispetto a quanto riscontrato nella variabile professione: i valori trovati sull'età sono consoni alla frequenza della professione studente; il 65% della popolazione (dal valore del quantile 10% di 18 al 3 quartile di 25) ha una età coerente con il valore 62% professione studente. La seconda è relativa alla forma della distribuzione: una mediana più bassa rispetto alla media indica una distribuzione obliqua a destra, con una concentrazione maggiore per i valori più bassi dell'intervallo, come è possibile vedere nel grafico. Questo è dimostrato anche dal valore dell'indicatore Skewness: il valore di 1,429 indica una distribuzione asimmetrica positiva. L'indicatore Kurtosis, con valore 3,837, indica una distribuzione ipernormale con una densità di frequenza elevata per i valori che si discostano maggiormente dalla media. Questi dati indicano come la popolazione dei giocatori sia concentrata maggiormente nell'età adolescenziale e giovanile, poiché il 90% della popolazione ha meno di 30 anni.

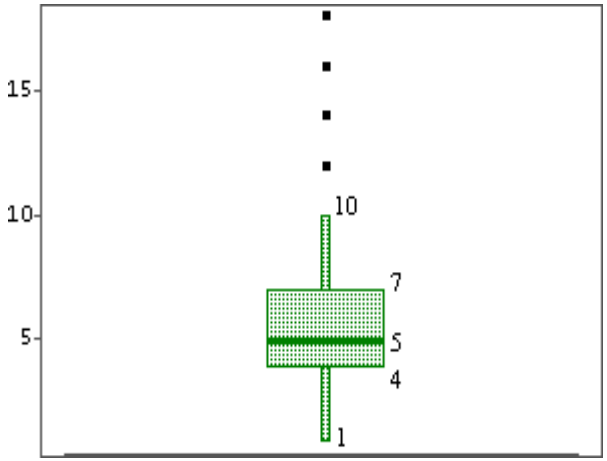
Questi dati, così come quelli relativi alla professione, indicano quindi una possibile utenza giovane e ancora nel periodo di studi.

Ore libere al giorno

Ore libere al giorno			
Momenti			
N	246	Somma dei pesi	246
Media	5,82114	Somma delle osservazioni	1.432
Deviazione std	2,56358	Varianza	6,57196
Skewness	1,25139	Kurtosis	2,62433
SS non corretta	9.946	SS corretta	1.610,13008
Coeff variaz	44,03921	Errore std media	0,16345

Misure statistiche di base			
Posizione		Variabilità	
Media	5,82114	Deviazione std	2,56358
Mediana	5	Varianza	6,57196
Moda	5	Intervallo	17
		Intervallo interquartile	3

Quantili	
Quantile	Stima
100% Max	18
99%	14
95%	10
90%	10
75% Q3	7
50% Mediana	5
25% Q1	4
10%	3
5%	3
1%	2
0% Min	1



Una volta individuato il genere, la professione e l'età del target, è utile capire quanto tempo libero essi abbiano al giorno, mediante l'utilizzo della variabile ore libere al giorno.

Avendo trovato una coerenza tra professione ed età, è logico aspettarsi anche una coerenza rispetto alle ore di tempo libero: infatti le misure statistiche di base di posizione (media, mediana e moda) indicano una certa congruità rispetto alle variabili precedenti. E' più probabile che uno studente abbia maggiore tempo libero rispetto a un impiegato o a un libero professionista, rispetto quindi a chi lavora.

In questo caso l'intervallo è di 17 ore (dato, però, da valori estremi come 18 e 14, derivante dalla piccola percentuale dei disoccupati), mentre l'intervallo interquartile è di 3 ore. Come nel caso precedente la forma della distribuzione è obliqua a destra, poiché la mediana 5 è inferiore alla media 5,821. Però, mentre l'indicatore Skewness 1,251 indica sempre una distribuzione asimmetrica positiva, l'indicatore Kurtosis, con valore 2,624, indica una distribuzione iponormale, con una ha densità di frequenza minore per valori molto distanti dalla media. Questo significa che si ha una maggiore distribuzione di frequenza per valori vicino alla media, ed è quindi importante tenere in considerazione che vi è più probabilità che il campione abbia come ore di tempo libero valori vicini alla media, come mostra il boxplot. Infatti la maggior parte delle osservazioni si concentra tra 4 e 7.

Tempo dedicato al gaming

Quantili Ore a casa Settimana	
Quantile	**Stima**
100% Max	130
99%	105
95%	70
90%	52
75% Q3	40
50% Mediana	27
25% Q1	17
10%	10
5%	6
1%	3
0% Min	2

Quantili Ore online a settimana	
Quantile	**Stima**
100% Max	100
99%	80
95%	35
90%	25
75% Q3	15
50% Mediana	10
25% Q1	5
10%	2
5%	1
1%	1
0% Min	0

Quantili % del tempo libero dedicato al gaming	
Quantile	**Stima**
100% Max	90
99%	90
95%	67
90%	50
75% Q3	40
50% Mediana	30
25% Q1	20
10%	10
5%	10
1%	5
0% Min	0

Coefficienti di correlazione di Pearson				
	Ore libere al giorno	**% tempo libero dedicato al gaming**	**Ore a casa Settimana**	**Ore online a Settimana**
Ore libere al giorno	1,00000 246	0,07145 0,2662 244	0,34092 <,0001 245	0,27088 <,0001 246
% tempo libero dedicato al gaming	0,07145 0,2662 244	1,00000 244	0,22088 0,0005 244	0,50776 <,0001 244
Ore a casa Settimana	0,34092 <,0001 245	0,22088 0,0005 244	1,00000 245	0,53768 <,0001 245
Ore online a Settimana	0,27088 <,0001 246	0,50776 <,0001 244	0,53768 <,0001 245	1,00000 246

Fondamentale per la nostra analisi è considerare il tempo che viene passato a casa settimanalmente, e in particolare on-line.

Utilizzando principalmente i quartili possiamo vedere come metà della popolazione trascorra il proprio tempo libero a casa settimanalmente per valori compresi tra 17 e 40 ore e che passi settimanalmente dalle 5 alle 15 ore online. Questi due dati sono molto importanti, perché ci indicano da un lato quanto tempo attualmente sono disposti a spendere su un videogames e dall'altro invece il potenziale di ore che potrebbero passare online. Infatti questo minor tempo passato online rispetto al tempo trascorso in casa è coerente con il momento di calma nel mercato: non vi sono giochi in grado di suscitare un grande interesse nei giocatori, così da invogliarli a passare un maggior tempo online. E' sicuramente un buon momento per creare un nuovo servizio con giochi in grado di creare maggior interesse.

Per effettuare un controllo su queste due variabili è stata utilizzata una terza variabile, la % del tempo libero dedicato al gaming. Sempre utilizzando i quartili si riscontra una certa coerenza, con il primo quartile che indica 20% di tempo libero dedicato al gaming on-line, la mediana con 30% e il 3 quartile con 40%. Dato che non esiste una univoca corrispondenza (infatti le percentuali si avvicinano a quelle reali, ma non sono precise) si è deciso di effettuare una analisi di correlazione tra le variabili di ore libere al giorno, ore a casa settimana, ore online a settimana e % del tempo dedicato al gaming tramite i coefficienti di correlazione di Pearson.

Non esiste una correlazione tra % tempo dedicato al gaming e ore libere al giorno; esiste però una correlazione tra il resto delle variabili: infati i valori 0,0005 e <0,0001 indicano come ore a casa settimana, ore online a settimana, % tempo libero dedicato al gaming siano effettivamente correlate tra loro, così come ore libere al giorno, ore a casa settimana e ore online a settimana.

Momenti a maggior afflusso

Misure statistiche di base Mattino			
Posizione		**Variabilità**	
Media	8,590	**Deviazione std**	19,462
Mediana	-	**Varianza**	378,761
Moda	-	**Intervallo**	100
		Intervallo interquartile	10
Skewness	3,016	**Kurtosis**	9,405

Misure statistiche di base Pomeriggio			
Posizione		**Variabilità**	
Media	30,119	**Deviazione std**	25,928
Mediana	25	**Varianza**	672,237
Moda	-	**Intervallo**	100
		Intervallo interquartile	40
Skewness	0,810	**Kurtosis**	0,072

Misure statistiche di base Sera			
Posizione		**Variabilità**	
Media	47,975	**Deviazione std**	28,021
Mediana	50	**Varianza**	785,201
Moda	50	**Intervallo**	100
		Intervallo interquartile	40
Skewness	-0,005	**Kurtosis**	-0,808

Misure statistiche di base Notte			
Posizione		**Variabilità**	
Media	13,365	**Deviazione std**	18,289
Mediana	5	**Varianza**	334,471
Moda	-	**Intervallo**	100
		Intervallo interquartile	20
Skewness	1,76	**Kurtosis**	3,467

Ultimo step dell'analisi del target di riferimento è capire in quale momento della giornata vi è maggior traffico di utenza.

Analizzando principalmente le misure statistiche di base si può notare come sia indiscutibile che sia la sera il momento in cui vi è maggior concentrazione di utenza. Questo è dato dal fatto che mediana e moda hanno valore 50% e la media si avvicina al valore 48%. Ciò significa che il 50% del tempo dedicato al gaming avviene di sera. Gli indicatori di distribuzione indicano inoltre che la frequenza è maggiormente concentrata presso la media. (lo Skewness indica una distribuzione quasi normale, mentre il valore negativo del Kurtosis indica una fortissima distribuzione di frequenza vicino alla media)

Come prevedibile però dai dati sull'età e sulla professione, il secondo momento della giornata è il pomeriggio, con una media di 30,12 e mediana di 25.

L'intervallo interquantile pari a 40% per entrambi i momenti ci forniscono una indicazione su come effettivamente sia concentrata maggiormente l'utenza in questi due momenti.

Caratteristiche Target

Queste sono le considerazioni ottenute dall'analisi sul target

- Possibilità di creazione del servizio.
- Coerenza delle variabili di analisi del target.
- Indicazioni su: sesso, età e professione del target per avere un quadro di riferimento sulla possibile utenza e per poter intraprendere strategie mirate.
- Indicazioni su tempo libero e su tempo dedicato al gaming, necessario per capire quanto un gioco deve essere impegnativo al fine di ottenere il loro interesse (non è possibile offrire un gioco che necessiti, per essere giocato, più del tempo che essi dedicano o hanno disponibile a settimana).
- Individuazione dei momenti a maggior afflusso di utenza sui server di gioco, al fine di fornire un maggior supporto nei momenti di massima affluenza, come la sera, senza però trascurare il pomeriggio.

Caratteristiche Core Business

Piattaforma di gioco				
	Frequenza	Percentuale	Frequenza cumulata	Percentuale Cumulata
Pc e Console	58	23,58	58	23,58
Esclusivamente PC	188	76,42	246	100,00

% del tempo passato a giocare in modalità multiplayer su console							
Momenti				Misure statistiche di base			
N	57	Somma dei pesi	57	Posizione		Quantili	
Media	24,68421	Somma osservazioni	1407	Media	24,68421	25% Q1	10
Deviazione std	23,71118	Varianza	562,21993	Mediana	20	50% Mediana	20
Skewness	1,61174	Kurtosis	2,15558	Moda	20	75% Q3	30

Numero di giochi							
Momenti				Misure statistiche di base			
N	246	Somma dei pesi	246	Posizione		Variabilità	
Media	2,83740	Somma delle osservazioni	698	Media	2,83740	Intervallo	12
Deviazione std	2,11457	Varianza	4,47141	Mediana	2	Intervallo interquartile	3
Skewness	1,35417	Kurtosis	2,70188	Moda	2	Osservazioni Estreme	5

Il primo step per individuare le caratteristiche del core business è capire se il servizio debba offrire server di gioco sia per PC che per Console. Solo il 23,58 degli intervistati ha dichiarato di utilizzare anche la Console per giocare on-line. Utilizzando la distribuzione di frequenza relativa esclusivamente ai giocatori che utilizzano PC e Console si nota come mediamente venga utilizzata solo per il 24,68% la Console, rispetto quindi a un 75,32% di utilizzo del PC. Inoltre lo Skewness mostra come la distribuzione sia anche asimmetrica positiva, e quindi obliqua a destra. Questo significa che l'utilizzo della console, oltre ad avere un valore relativamente basso di circa 25%, ha una distribuzione di frequenza concentrata per valori minori rispetto alla media. Infine i quartili indicano come il 50% di coloro che utilizzano la Console giocano online con essa per valori compresi tra il 10% e il 30%. Quindi, se circa il 25% della popolazione utilizza la Console online per solo circa 25% è ragionevole escludere un servizio che offra anche server di gioco per console. Infatti analizzando la variabile della % dell'utilizzo del PC per giocare in rete (qui non riportato perché non rilevante all'analisi tra pc e console) si ottengono valori di media mediana e moda compresi tra il 90 e il 99%.

Individuata, quindi, la piattaforma di gioco è possibile procedere all'analisi del numero di giochi offerti e della tipologia. L'analisi sul numero dei giochi a cui l'intervistato sta attualmente giocando mostra una media di quasi 3 giochi, ma una moda di 2. La distribuzione è asimmetrica positiva con concentrazione di frequenza di distribuzione per valori vicino alla media. L'intervallo interquartile è di 3 giochi. E' possibile quindi affermare che ogni giocatore è portato a giocare da 1 a 3 giochi.

I generi su cui puntare

Genere principale				
Genere	Frequenza	Percentuale	Frequenza cumulata	Percentuale cumulata
MMO	122	49,59	122	49,59
FPS	86	34,96	208	84,55
Browser Game	12	4,88	220	89,43
RTS	12	4,88	232	94,31
Social	7	2,85	239	97,16
Sport	6	2,44	245	99,6
Carte	1	0,40	246	100

Genere secondario				
Genere	Frequenza	Percentuale	Frequenza cumulata	Percentuale cumulata
FPS	86	34,96	86	34,96
MMO	53	21,54	139	56,50
RTS	35	14,23	174	70,73
Sport	28	11,38	202	82,11
Browser Game	21	8,54	223	90,65
Social	13	5,28	236	95,93
Carte	5	2,04	241	97,97
Missing	5	2,04	246	100

E' stato chiesto agli intervistati di fornire una doppia preferenza, una sul genere principale e una sul genere secondario, in modo da fornire una offerta di giochi capace di soddisfare le loro richieste in coerenza al numero di giochi giocati (moda 2).

Dall'analisi del genere principale risulta che il 50% degli intervistati predilige il genere MMO (Massive Multyplayer On-line), seguito da un 35% che predilige invece gli FPS (First Person Shooter): raggiungiamo così l'85% della preferenza degli utenti per il genere principale.

Dall'analisi del genere secondario FPS e MMO coprono il 56%, il genere RTS (Real Time Strategy) ha una preferenza del 15%.

L'offerta deve quindi contenere sicuramente MMO e FPS, che sono i due generi pilastro sia come preferenza principale, sia come preferenza secondaria.

Poiché il genere FPS richiede poche risorse in termini di tempo per la gestione, è possibile aggiungere un terzo genere, come l'RTS, anche esso poco dispendioso in termini di tempo.

Infine è utile inserire un quarto genere in modo da coprire il 95% della preferenza principale e quasi l'80% della preferenza secondaria: browser game. Essi infatti richiedono un maggior tempo di gestione da parte del servizio (non superiore al MMO, ma molto inferiore rispetto all'RTS e all'FPS), ma solitamente un minor tempo di impegno da parte dell'utente: esso infatti viene utilizzato su browser e molte volte viene giocato in contemporanea, tenendolo in background, a un altro genere.

L'analisi di correlazione tra generi ci fornirà inoltre una buona indicazione sulla preferenza del genere secondario in base alla preferenza del genere principale.

Correlazione tra generi

Tabella di correlazione tra genere pincipale e genere secondario								
Genere principale	Genere secondario							Totale
	Browser Game	Carte	FPS	MMO	RTS	Social	Sport	
Browser Game	-	-	3	5	-	1	2	11
	-	-	1,24	2,07	-	0,41	8,83	4,56
	-	-	27,27	45,45	-	9,09	18,18	
	-	-	3,49	9,43	-	7,69	7,14	
Carte	1	-	-	-	-	-	-	1
	0,41	-	-	-	-	-	-	0,41
	100,00	-	-	-	-	-	-	
	4,76	-	-	-	-	-	-	
FPS	7	1	-	41	18	3	14	84
	2,90	0,41	-	17,01	7,47	1,24	5,81	34,85
	8,33	1,19	-	48,81	21,43	3,57	16,67	
	33,33	20,00	-	77,36	51,43	23,08	50,00	
MMO	10	3	74	-	15	9	9	120
	4,15	1,24	30,71	-	6,22	3,73	3,73	49,79
	8,33	2,50	61,67	-	12,50	7,50	7,50	
	47,62	60,00	86,05	-	42,86	69,23	32,14	
RTS	-	-	5	5	-	-	2	12
	-	-	2,07	2,07	-	-	0,83	4,98
	-	-	41,67	41,67	-	-	16,67	
	-	-	5,81	9,43	-	-	7,14	
Social	2	1	1	1	1	-	1	7
	0,83	0,41	0,41	0,41	0,41	-	0,41	2,90
	28,57	14,29	14,29	14,29	14,29	-	14,29	
	9,52	20,00	1,16	1,89	2,86	-	3,57	
Sport	1	-	3	1	1	-	-	6
	0,41	-	1,24	0,41	0,41	-	-	2,49
	16,97	-	50,00	16,67	16,67	-	-	
	4,76	-	3,49	1,89	2,86	-	-	
Totale	21	5	86	53	35	13	28	241
	8,71	8,71	2,07	35,68	21,99	14,52	5,39	100

Frequenza dei valori mancanti = 5

	DF	Valore	Prob		Valore	
Chi-quadrato	36	150	<.0001	Coefficiente Phi	0,789	ATTENZIONE: 78% delle celle ha conteggi previsti inferiori a 5. Il chi-quadrato potrebbe non essere un test valido.
Chi-quadrato rapp verosim	36	192	<.0001	Coefficiente di contingenza	0,6194	
Chi-quadrato MH	1	9	0.0021	V di Cramer	0,3221	

La tabella di correlazione tra genere principale e genere secondario mostra come vi sia una forte correlazione tra MMO e FPS. Il 61,67 di chi ha scelto come genere principale MMO ha scelto come secondario FPS, mentre il 48,81 di chi ha scelto come genere principale FPS ha scelto come secondario MMO: questo indica come questi due siano i generi principali e fondamentali del servizio.

Così come chi ha scelto Browser game o RTS come principale ha preferito maggiormente FPS o MMO come secondario, mentre per chi ha scelto MMO o FPS come principale abbiamo buone % di preferenza su Browser game e RTS.

Sommando il numero di osservazioni trovate nell'incrocio tra questi 4 generi, tra principale e secondario, troviamo un valore di 183 osservazioni, pari al 75% degli intervistati. Questo dato mostra come questi 4 generi possano portare molti giocatori a usufruire del servizio.

Infine l'indice Chi-quadro mostra come queste 2 variabili siano correlate tra loro, anche se bisogna notare che il 78% delle celle ha conteggi previsti inferiori a 5, e che quindi il chi-quadro potrebbe non essere un test valido. Comunque bisogna ammettere che dalla tabella esiste un certo riscontro sulla correlazione dei generi (oltre a FPS e MMO, anche Browser game e RTS mostrano una forte correlazione, avendo sempre 0 come valore nei loro incroci).

Modello di business per genere

Browser Game				
Modello di Business	Frequenza	Percentuale	Frequenza cumulata	Percentuale cumulata
Free	9	75,00	9	75,00
Free + Acconto	1	8,33	10	83,33
Free + acconto	1	8,33	11	91,66
Gioco	1	8,33	12	100,00

FPS				
Modello di Business	Frequenza	Percentuale	Frequenza cumulata	Percentuale cumulata
Gioco	58	67,44	58	67,44
Free	19	22,09	77	89,53
Abbonamento	5	5,81	82	95,34
Free + Oggetti	2	2,33	84	97,67
Free + acconto	2	2,33	86	100

MMO				
Modello di Business	Frequenza	Percentuale	Frequenza cumulata	Percentuale cumulata
Abbonamento	46	37,70	46	37,7
Free	37	30,33	83	68,03
Gioco	31	25,41	114	93,44
Free + Oggetti	6	4,92	120	98,36
Free + Acconto	2	1,64	122	100

RTS				
Modello di Business	Frequenza	Percentuale	Frequenza cumulata	Percentuale cumulata
Free	4	33,33	4	33,33
Gioco	8	66,67	12	100

Individuati i quattro generi, è utile capire quale sistema di business adottare in base alle preferenze espresse, utilizzando una divisione per classi della variabile modalità di business tramite la variabile genere principale (viene utilizzata solo la variabile genere principale poiché la secondaria si scosta di poco da questi valori).

Per i generi FPS e RTS la preferenza è quasi esclusiva, con un 67% sulla modalità di acquisto del videogioco e di usufrutto dei server di gioco senza pagare. E' possibile adottare questa modalità di business, poiché non richiedono grandi sforzi di gestione questi due generi. La creazione di server dedicati a questi due generi può favorire il servizio da un lato per diffondere il nome all'interno delle diverse comunità di giocatori e dall'altro per informare i diversi giocatori sulle future uscite di offerte relativi ad altri generi. Inoltre è possibile creare un ritorno economico con attività di co-marketing con i distributori di videogames, che verrà spiegata nella parte sulle caratteristiche dei servizi accessori.

Per il genere Browser Game invece il 75% di coloro che preferiscono il genere vorrebbero un sistema free. Anche in questo caso è possibile utilizzare questa modalità di business poiché trattandosi di giochi su browser è possibile creare direttamente sulla pagina di gioco dei banner pubblicitari.

Infine per il genere MMO la preferenza è per 37,70% abbonamento, 30,33% Free e 25% acquisto del gioco. Utilizziamo quindi la tabella con i modelli di business utilizzati attualmente per i MMO: il 52,04% utilizza un sistema ad abbonamento, il 33,68% free e solo il 7,14 il modello ad acquisto del gioco. E' quindi possibile pensare a un modello di business che offra 30 giorni di prova gratuita e successivamente il pagamento di un abbonamento mensile.

MMO				
Modello di Business	Frequenza	Percentuale	Frequenza Cumulata	Percentuale cumulata
Abbonamento	51	52,04	51	52,04
Free	33	33,68	84	85,72
Free + Oggetti	7	7,14	91	92,86
Gioco	7	7,14	98	100

Prezzo abbonamento

Misure statistiche di base			
Posizione		Variabilità	
Media	12,71	Deviazione std	4,06
Mediana	13	Varianza	16,46
Moda	15	Intervallo	17,00
Skewness	-0,5	Intervallo interquartile	4,00

Misure statistiche di base			
Posizione		Variabilità	
Media	13,31	Deviazione std	3,845
Mediana	13	Varianza	14,786
Moda	12	Intervallo	24
Skewness	0,2	Intervallo interquartile	3

Disponibilità	
Quantile	Stima
100% Max	20
99%	20
95%	20
90%	16
75% Q3	15
50% Mediana	13
25% Q1	11
10%	6
5%	5
1%	3
0% Min	3

Spesa reale	
Quantile	Stima
100% Max	24
99%	24
95%	20
90%	20
75% Q3	15
50% Mediana	13
25% Q1	12
10%	10
5%	10
1%	0
0% Min	0

L'ultimo step dell'analisi delle caratteristiche del core business, dopo aver individuato piattaforma, generi e sistema di business per i generi offerti è capire, per l'unico genere che prevede un esborso per l'utente, quanto debba essere il valore di un abbonamento mensile.

Analizzando la disponibilità a spendere e la spesa reale per un abbonamento mensile si nota un intervallo compreso tra 11 e 15 € per la disponibilità e un intervallo compreso tra 12 e 15€ per la spesa reale. Da un lato però si ha una disponibilità media a spendere di meno (12,71€) contro la media di spesa reale (13,31€). La moda indica che il prezzo maggiormente praticato è di 12€ per un abbonamento. L'indicatore di Skewness ci indica che la distribuzione per la disponibilità è asimmetrica negativa, mentre la distribuzione per la spesa reale è asimmetrica positiva. Questo significa che i giocatori attualmente pagano maggiormente per valori inferiori a 13€, ma che sarebbero disposti a pagare anche di più.

E' possibile quindi posizionare una forbice di prezzo per abbonamento compreso tra 11 e 15 euro, in base al MMO scelto.

Caratteristiche Core Business

Dall’analisi delle variabili sul core business risulta che:

- Il servizio deve puntare esclusivamente su server di gioco per PC.
- Il servizio deve tenere in considerazione che ogni utente gioca mediamente a 2 / 3 giochi.
- I generi principali individuati sono MMO, FPS, RTS, e Browser game
- Le modalità di business per i diversi generi sono: acquisto del gioco per FPS e RTS, Free per il Browser e Abbonamento da 11 a 15€ per MMO con possibilità di 30 giorni di prova Free.
- Attività di co-marketing su Browser game, FPS e RTS

Servizi accessori

Chat vocale				
Utilizzo:	Frequenza	Percentuale	Frequenza cumulata	Percentuale cumulata
No	44	17,89	44	17,89
Si	202	82,11	246	100

Forum				
Utilizzo:	Frequenza	Percentuale	Frequenza cumulata	Percentuale cumulata
No	3	1,22	3	1,22
Si	243	98,78	246	100

Esempio categoria ad alto interesse				
Livello di preferenza	Frequenza	Percentuale	Frequenza Cumulata	Percentuale cumulata
.	1	0,41	1	0,41
5	3	1,22	8	3,25
10	22	8,94	43	17,48
15	59	23,98	113	45,93
20	70	28,46	246	100

Esempio di categoria a medio interesse				
Livello di preferenza	Frequenza	Percentuale	Frequenza cumulata	Percentuale cumulata
.	1	0,41	1	0,41
5	30	12,20	95	38,62
10	35	14,23	153	62,20
15	24	9,76	198	80,49
20	19	7,72	246	100

Un primo servizio accessorio è la disponibilità di un server di chat vocale: l'82,11 ha dichiarato che utilizza o utilizzerebbe il servizio.

Un secondo servizio è invece quello di spazi dedicati a discussione come il forum: la quasi totalità degli intervistati ha dichiarato di utilizzare o che utilizzerebbe un forum. La struttura dei forum dipende, però, dal grado di interesse alle diverse categorie. Sono state individuate 3 tipologie di categorie: categorie ad alto interesse, con 50% delle osservazioni nei livelli di preferenza tra 15 e 20, categorie a medio interesse, con il 50% delle osservazioni nei livelli di preferenza tra 10 e 20, categorie a basso interesse, le restanti.

Sono 4 le categorie ad alto interesse: cinema, multiplayer, musica e PC. Per esse verranno create 4 sezioni separate.

Le categorie a medio interesse sono invece 4: Console, fumetti, singleplayer, e telefilm. Console e singleplayer verranno raggruppate insieme in una categoria specifica, telefilm verrà inserita insieme alla categoria cinema e fumetti spostata nella sezione generale.

Le restanti categorie verranno inserite nella sezione generale

Compravendita di oggetti

Compravendita				
Favorevole:	Frequenza	Percentuale	Frequenza Cumulata	Percentuale cumulata
No	175	71,14	175	71,14
Si	71	28,86	246	100

Gestore compravendita				
Soggetto:	Frequenza	Percentuale	Frequenza cumulata	Percentuale cumulata
Giocatori	52	21,14	52	21,14
Servizio	194	78,86	246	100

Relazione sulla compravendita			
	Soggetto		
Favorevole	Giocatori	Servizio	Totale
No	27 10,98 15,43 51,92	148 60,16 84,57 76,29	175 71,14
Si	25 10,16 35,21 48,08	46 18,7 64,79 23,71	71 28,86
Totale	52 21,14	194 78,86	246 100

Statistica	DF	Valore	Prob
Chi-quadrato	1	11,858	0,0006
Chi-quadrato rapp verosim	1	11,118	0,0009
Chi-quadrato corr continuità	1	10,701	0,0011
Chi-quadrato MH	1	11,809	0,0006
Coefficiente Phi		-0,2195	
Coefficiente di contingenza		0,2144	
V di Cramer		-0,2195	

Un ulteriore servizio è relativo alla compravendita di oggetti all'interno dei diversi giochi utilizzando denaro reale. Quasi il 72% degli intervistati ha dichiarato di non essere favorevole.

In caso, però, vi fosse una maggiore richiesta di questo servizio in futuro l'80% ha dichiarato che fosse il servizio di gaming stesso a gestire le transazioni. Questo si verificherebbe anche in futuro se coloro che oggi hanno detto di non essere a favore lo diventerebbero: infatti il test chi-quadrato, con un valore p-value basso, indica che esiste una correlazione tra la volontà di avere il servizio e chi lo debba gestire. Anche nella tabella della relazione sulla compravendita è possibile riscontrare questa tendenza.

Servizio di acquisto on-line

Possibilità di scaricare i giochi dal sito				
	Frequenza	Percentuale	Frequenza cumulata	Percentuale cumulata
No	90	36,59	90	36,59
Si	156	63,41	246	100

Utilizzo della modalità singleplayer				
	Frequenza	Percentuale	Frequenza cumulata	Percentuale cumulata
No	22	8,94	22	8,94
Si	224	91,06	246	100

% utilizzo del pc per la modalità singleplayer			
Posizione		Variabilità	
Media	75,830	Deviazione std	32
Mediana	99	Varianza	1048
Moda	100	Intervallo	100
		Intervallo interquartile	50

Quantile	Stima
100% Max	100
99%	100
95%	100
90%	100
75% Q3	100
50% Mediana	99
25% Q1	50
10%	20
5%	10
1%	0
0% Min	0

Ultimo servizio, capace di creare oltre che una caratteristica distintiva da quelli esistenti, anche una fonte di ricavo, è lo shop on-line.

Questo servizio è relativo alla possibilità di acquistare i videogiochi direttamente dal sito, scaricandoli in formato ISO, tramite la creazione di un acconto che ricordi ogni singolo gioco acquistato. Il 63,41% si è dichiarato favorevole all'introduzione di questo servizio, mentre il 36,59% ha dichiarato una preferenza per l'acquisto del gioco fisico on-line o in negozio. Il servizio quindi potrebbe, per coprire una percentuale di acquirenti più alta del 65% circa, di creare uno shop sia delle copie in formato ISO, sia dei giochi fisici tramite spedizione.

Grazie a questo servizio è possibile inoltre effettuare una strategia di co-marketing con i distributori di videogames: infatti il servizio oltre a mettere in vendita i videogames, può creare per i 2 generi FPS e RTS dei server dedicati, in modo che al lancio del nuovo gioco del diverso genere vi siano già server disponibili dove provarli in multiplayer.

Infine l'ultima domanda riguarda la vendita di giochi esclusivamente per PC o anche per Console: il 91,06% degli intervistati ha dichiarato di giocare anche in modalità singleplayer e in particolare utilizzano mediamente per 75,83% il PC. Anche dall'intervallo interquartile risulta un utilizzo tra il 50 e il 100%, per cui è possibile trascurare anche in questo caso le console.

Servizi Accessori

Dall’analisi delle variabili sui servizi accessori risulta:

- Creazione di un servizio di chatvocale
- Creazione di un forum con 7 sezioni in base ai diversi argomenti: Pc e singleplayer, Multiplayer, Console e singleplayer, Musica, Cinema e TV, Generale, Spam (in modo da contenere discussioni inutili).
- Attuale congelamento del servizio di compravendita di oggetti tramite denaro reale. In caso diventasse un esigenza, la possibilità di gestione sarebbe nelle mani del servizio.
- Creazione di uno shop on-line di videogiochi per PC, sia in formato fisico con scatola, sia in formato elettronico con ISO.
- Attività di co-marketing per la creazione di server dedicati per gli FPS e RTS (in modo da generare un profitto con le vendite dei giochi).

Modalità di pagamento

Postepay				
Livello di preferenza	Frequenza	Percentuale	Frequenza cumulata	Percentuale cumulata
.	20	8,13	20	8,13
1	46	18,70	66	26,83
2	3	1,22	69	28,05
3	5	2,03	74	30,08
5	14	5,69	88	35,77
6	12	4,88	100	40,65
7	13	5,28	113	45,93
8	20	8,13	133	54,06
9	113	45,94	246	100

Bollettino				
Livello di preferenza	Frequenza	Percentuale	Frequenza cumulata	Percentuale cumulata
.	20	8,13	20	8,13
1	164	66,67	184	74,80
2	8	3,25	192	78,05
3	12	4,88	204	82,93
4	8	3,25	212	86,18
5	16	6,50	228	92,68
6	7	2,85	235	95,53
7	4	1,63	239	97,16
8	4	1,63	243	98,79
9	3	1,21	246	100

Ultima analisi, prima di passare a quelle basate su modelli, è relativa alle modalità di pagamento, sia per l'abbonamento, sia per i videogiochi.

Sono 5 le modalità individuate: Postepay, Carta prepagata, Carta di Credito, Bollettino postale e Bonifico bancario.

La modalità che ha un livello di preferenza maggiore è la Postepay, con un 45,94 di intervistati che indica su scala da 1 a 9 il massimo. Segue carta prepagata, con un livello medio di 6/7. Terza la Carta di credito che mostra un livello di preferenza 1 pari al 40% e un 50% per valori invece alti. Bassissimi livelli di preferenza si hanno per Bollettino postale e Bonifico bancario.

Bisogna quindi prevedere come modalità di pagamento Postepay, Cartaprepagata, Carta di credito. Eventualmente può essere inserito il bollettino postale in caso qualche utente non disponga di carte prepagate o di credito.

Analisi Fattoriale

Il modello di analisi fattoriale viene utilizzato per spiegare un fenomeno attraverso l'uso di fattori, che riducono il numero di variabili utilizzate, poiché sintetizzano l'informazione contenuta nelle diverse variabili.

Sono state individuate, tramite l'analisi di correlazione, 25 variabili fondamentali che spiegano il comportamento generale di un giocatore: free_ore_giorno_4, videogiochi_5, ore_casa_6, ore_online_7, PCm_10, ngiochi_11, durata_15, durata_21, spesa_27, cartacredito_28, cartaprepagata_28, postepay_28, paypal_28, bollettino_28, bonifico_28, grafica_31, ambientazione_31, storia_31, scopo_31, giocabilita_31, pvp_31, interazione_31, online_38, single_38, s_pc_40.

Da questo set di variabili verrà individuato il numero ottimale di fattori che spiegano il fenomeno preso in analisi e successivamente verranno interpretati.

Scelta del numero dei fattori

Autovalori della matrice di correlazione: Totale = 25 Media = 1				
	Autovalore	Differenza	Proporzione	Cumulata
1	3.08934203	0.54911095	0.1236	0.1236
2	2.54023109	0.46559439	0.1016	0.2252
3	2.07463670	0.39476035	0.0830	0.3082
4	1.67987635	0.07366017	0.0672	0.3754
5	1.60621618	0.19633990	0.0642	0.4396
6	1.40987628	0.18605759	0.0564	0.4960
7	1.22381869	0.01929207	0.0490	0.5450
8	1.20452662	0.07676370	0.0482	0.5931
9	1.12776292	0.10908435	0.0451	0.6383
10	1.01867856	0.08006948	0.0407	0.6790
11	0.93860908	0.07833389	0.0375	0.7165
12	0.86027520	0.08612879	0.0344	0.7510
13	0.77414641	0.03519307	0.0310	0.7819
14	0.73895334	0.04474980	0.0296	0.8115
15	0.69420354	0.04042425	0.0278	0.8392
16	0.65377929	0.05237684	0.0262	0.8654
17	0.60140246	0.05202639	0.0241	0.8895
18	0.54937606	0.02993691	0.0220	0.9114
19	0.51943916	0.08922700	0.0208	0.9322
20	0.43021216	0.03855207	0.0172	0.9494
21	0.39166009	0.06688664	0.0157	0.9651
22	0.32477345	0.04689385	0.0130	0.9781
23	0.27787960	0.00755485	0.0111	0.9892
24	0.27032475	0.27032475	0.0108	1.0000
25	0.00000000		0.0000	1.0000

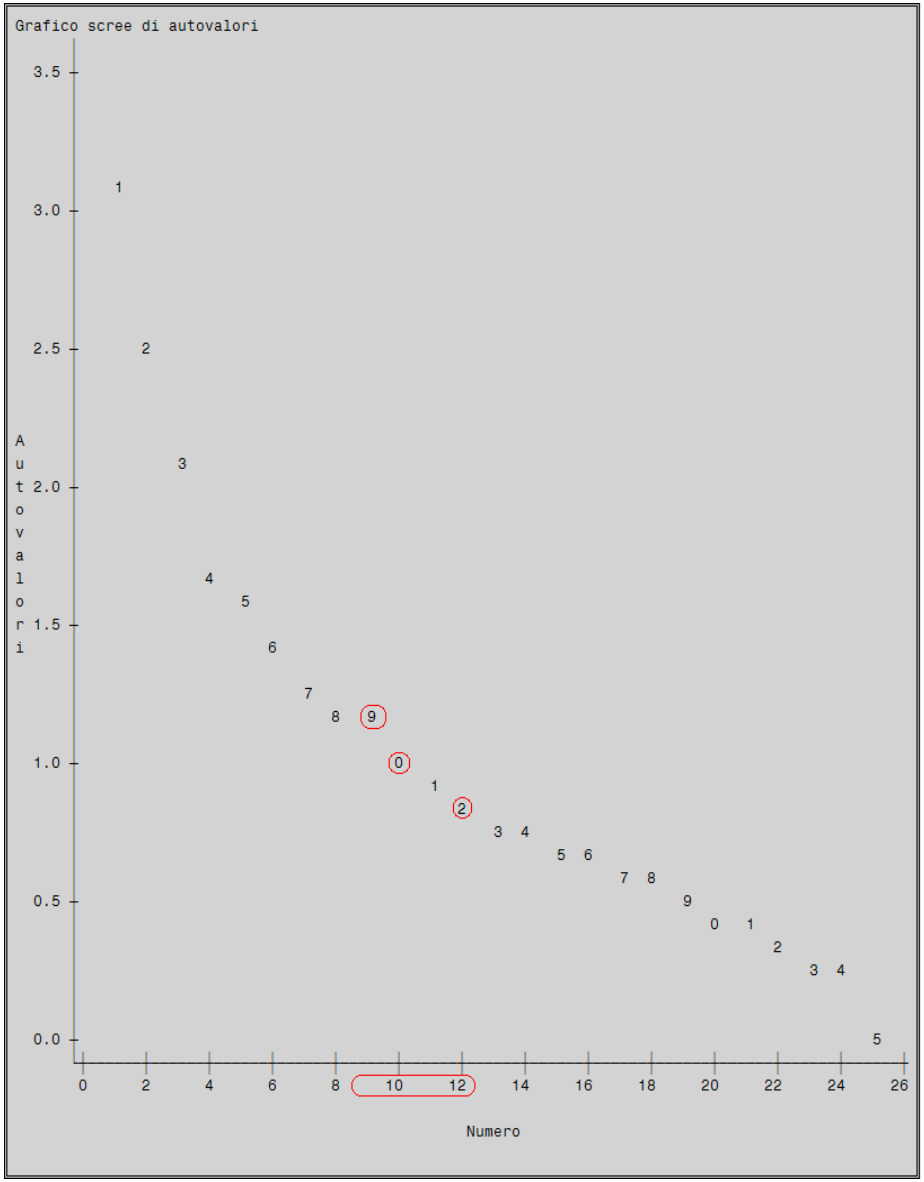
Il primo step dell'analisi fattoriale è individuare il numero ottimale di fattori che spiegano in maniera discreta il fenomeno.

Un primo indicatore può essere il rapporto tra numero di fattori e variabili: il numero dei fattori deve essere circa il 25% del numero delle variabili. Nel nostro caso i fattori dovrebbero essere quindi 6.

Un secondo indicatore è invece la percentuale di varianza spiegata: essa deve essere compresa tra il 60% e il 75%. Secondo questo indicatore 6 fattori spiegando una varianza del 50% non sono ottimali; bisogna prendere in considerazione quindi un numero di fattori compreso tra 9 e 12.

Per poter scegliere tra questo intervallo il numero ottimale di fattori utilizziamo quindi un terzo metodo: il metodo dell'autovalore (come suggerisce SAS) o criterio di Mineigen. Viene preso in considerazione il numero di fattori con autovalore maggiore di 1. Per cui il numero di fattori ottimali è 10.

Analisi dello Scree Plot



Analizzando lo scree plot risulta però che il decimo fattore cade in una zona obliqua della curva, mentre il nono e il dodicesimo fattore cadono all'inizio di un appiattimento della curva.

Non potendo decidere a priori quale sia il numero ottimale di fattori si procederà a una analisi a 9 fattori, a 10 fattori e a 12 fattori, tenendo non solo conto della variabilità spiegata dei singoli fattori, ma anche dell'interpretabilità e della necessità di avere il minor numero possibile di fattori per semplificare l'analisi.

L'analisi a 7 fattori è esclusa a priori per l'insufficienza di variabilità spiegata.

Bisogna aggiungere una considerazione finale: attraverso l'analisi dello scree plot per scegliere il numero di fattori risulterebbe che il numero ottimale di fattori sarebbe 8 o 13. Questi due valori però non possono essere presi in considerazione, poiché 8 fattori spiegano poco più del 50% della varianza totale, mentre 13 fattori spiegano più del 75%, ma sono in numero eccessivo (sono più della metà delle variabili originali)

Scelta del modello

Comunalità finale per ogni variabile			
	9 Fattori	10 Fattori	12 Fattori
free_ore_giorno_4	0.57233176	0.57878001	0.63578842
Videogiochi_5	0.44048673	0.46042205	0.83151974
ore_casa_6	0.53715040	0.58772484	0.75538083
ore_online_7	0.76720026	0.76907719	0.77614754
PCm_10	0.70066639	0.72053507	0.74589904
ngiochi_11	0.59073281	0.86111079	0.86799761
durata_15	0.68889660	0.70597372	0.75016860
durata_21	0.74852856	0.76380749	0.78649846
spesa_27	0.47842870	0.65015884	0.77663736
cartacredito_28	0.54902769	0.59522049	0.77261747
cartaprepagata_28	0.59128508	0.61469818	0.73934989
postepay_28	0.63238146	0.63815590	0.72584913
paypal_28	0.52839049	0.58481982	0.66603103
bollettino_28	0.75282753	0.75287182	0.76973607
bonifico_28	0.77931424	0.78150103	0.81088647
grafica_31	0.57025397	0.63603617	0.74873350
ambientazione_31	0.68053778	0.69493515	0.72089278
storia_31	0.62891964	0.62922787	0.66148245
scopo_31	0.52230325	0.56723001	0.66805627
giocabilita_31	0.63687487	0.64015109	0.64103536
pvp_31	0.55141167	0.56541971	0.57703745
interazione_31	0.49263502	0.56073066	0.72654852
online_38	0.89330905	0.94364520	0.94513629
single_38	0.89330905	0.94364520	0.94513629
s_pc_40	0.72908385	0.72908709	0.72928313

Tramite l'analisi delle comunalità finali nei diversi modelli a 9, 10 e 12 fattori risulta che la varianza di ciascuna delle 25 variabili è spiegata tutte al di sopra del 50% solo se si utilizzano 12 fattori. Quando si utilizzano 10 fattori troviamo 1 solo valore al di sotto dello 0,5 (anche se il valore è di 0,46) per la variabile videogiochi. Nell'analisi a 9 fattori invece sono 3 le variabili da tenere in considerazione: videogiochi con 0,44 di varianza spiegata, spesa con 0,47 e interazione con 0,49.

Tenendo conto, come detto in precedenza, della necessità di ridurre il numero delle variabili, della varianza spiegata e dell'interpretabilità si è scelto di adottare il modello a 9 fattori (varianza spiegata totale 63%), poiché la varianza di due variabili sono prossime allo 0,5 e la terza è correlata ad altre 2 variabili ben spiegate nel modello (si veda la correlazione tra videogiochi, ore a casa settimanali e ore online) e poiché esiste una interpetabilità migliore con 9 fattori.

Interpretazione

Schema fattoriale ruotato									
	Factor1	Factor2	Factor3	Factor4	Factor5	Factor6	Factor7	Factor8	Factor9
online_38	0.91611	.	Interazione		.	.	.	.	.
interazione_31	0.42152	.			.	.	.	.	.
pvp_31	0.38251	.			.	.	.	.	.
single_38	-0.91611	.			.	.	.	.	.
ore_online_7	.	0.80111	.	Tempo dedicato al gaming		.	.	.	.
ore_casa_6	.	0.72604	.			.	.	.	.
free_ore_giorno_4	.	0.6886	.			.	.	.	.
Videogiochi_5	.	0.41108	.			.	.	.	.
ambientazione_31	.	.	0.79885	.	Caratteristiche del gioco		.	.	.
storia_31	.	.	0.7602	.			.	.	.
grafica_31	.	.	0.65554	.			.	.	.
s_pc_40	.	.	.	0.84284	.	Utilizzo del PC		.	.
PCm_10	.	.	.	0.81539	.			.	.
bonifico_28	.	.	.	.	0.85884	.	Modalità di pagamento cartaceo		
bollettino_28	.	.	.	.	0.83825	.			
durata_21	.	.	Impegno in un gioco		.	0.84473	.	.	.
durata_15	.	.			.	0.68007	.	.	.
postepay_28	.	.	.	Modalità di pagamento on-line		.	0.71106	.	.
paypal_28	.	.	.			.	0.62874	.	.
cartaprepagata_28	.	.	.			.	0.62112	.	.
cartacredito_28	.	.	.	.	Proattività al gaming		.	0.71663	.
spesa_27	.	.	.	.			.	0.51216	.
ngiochi_11	.	.	.	.			.	0.48697	.
giocabilita_31	.	.	.	.	.	Motivazione a giocare		.	0.73975
scopo_31	.	.	.	.	.			.	0.49277

Regressione Lineare

Il modello di regressione lineare ha lo scopo di trovare una relazione tra una variabile dipendente e altre variabili denominate regressori; in altre parole si vuole capire un determinato fenomeno spiegato da una variabile attraverso l'utilizzo di altre variabili.

Nel nostro caso si è scelto di creare il modello sulla variabile dipendente Durata_15, che spiega la durata in termini di giorni di un possibile utente su un videogioco.

Questa scelta è dettata da due motivazioni: la prima è sicuramente collegata all'importanza di capire questo fenomeno, da cui è possibile individuare quali giocatori utilizzeranno maggiormente il servizio; la seconda è di carattere didattico, poiché anche includendo tutte le variabili nel modello solamente questa variabile ottiene un valore R-quadro attorno allo 0,5, mentre provando altri incroci il valore è sempre compreso tra 0,3 e 0,37. Già con un valore di 0,5 il modello non spiega in maniera esausitiva il fenomeno, con valori ancora inferiori era praticamente inutile studiare un modello per poi applicarlo.

Coefficienti di correlazione di Pearson			
	durata_15		durata_15
Eta_2 Eta_2	0.18738 0.0032 245	modellofreed_26 modellofreed_26	0.10783 0.0922 245
Videogiochi_5 Videogiochi_5	0.09909 0.1219 245	modelloabbonamentod_26 modelloabbonamentod_26	0.12238 0.0558 245
ore_online_7 ore_online_7	0.2108 0.0009 245	modellomistod_26 modellomistod_26	-0.15738 0.0137 245
solopc_8 solopc_8	0.21223 0.0008 245	grafica_31 grafica_31	-0.14352 0.0247 245
PCm_10 PCm_10	0.14936 0.0193 245	ambientazione_31 ambientazione_31	-0.13452 0.0354 245
Consolem_10 Consolem_10	-0.14503 0.0232 245	storia_31 storia_31	-0.13831 0.0304 245
ngiochi_11 ngiochi_11	-0.10731 0.0938 245	pvp_31 pvp_31	-0.12228 0.056 245
durata_21 durata_21	0.49172 <.0001 238	interazione_31 interazione_31	0.14206 0.0262 245
generefpsd_25 generefpsd_25	-0.15181 0.0174 245	perfomance_35 perfomance_35	-0.10429 0.1034 245
generebrowser_25 generebrowser_25	0.10317 0.1072 245	online_38 online_38	0.39168 <.0001 245
generemmod_25 generemmod_25	0.13889 0.0297 245	soddisfazione_42 soddisfazione_42	0.14593 0.0223 245

Scelta dei regressori

Per scegliere il set di regressori di partenza è stato utilizzato l'indice di correlazione tra la variabile dipendente e le altre variabili. Sebbene un buon regressore debba avere una coefficiente di correlazione maggiore di 0,4, nel nostro caso si è dovuto abbassare la scelta a 0,1. E' stato quindi individuato un set di partenza di 22 regressori.

Alcuni di questi regressori, come modellofreed_26 per esempio, sono il risultato della costruzione di variabili dummy per poter utilizzare anche variabili qualitative nel modello. Queste variabili sono caratterizzate dal codice finale d_numero della variabile qualitativa di partenza.

Modello a 22 Regressori

Stime dei parametri. R-quadro= 0,45 R-quadro corr= 0,4					
Variabile	Stima dei parametri	Errore standard	Valore t	Pr > |t|	Stima standardizzata
Interc.	-7146.27035	3525.6711	-2.03	0.0439	0
Eta_2	7.56154	4.61432	1.64	0.1027	0.08762
Videogiochi_5	0.62593	1.58125	0.4	0.6926	0.02449
ore_online_7	1.80382	2.16178	0.83	0.405	0.05589
solopc_8	93.79163	73.4589	1.28	0.2031	0.09164
PCm_10	70.07065	35.14696	1.99	0.0475	2.50477
Consolem_10	69.94577	35.0622	1.99	0.0473	2.51299
ngiochi_11	-23.13933	11.84966	-1.95	0.0521	-0.11252
durata_21	0.57082	0.0846	6.75	<.0001	0.40253
generefpsd_25	-108.2981	85.16752	-1.27	0.2049	-0.11546
generebrowser_25	-84.61392	101.43847	-0.83	0.4051	-0.0608
generemmod_25	-67.66981	83.77925	-0.81	0.4201	-0.07633
modellofreed_26	86.07122	93.47403	0.92	0.3582	0.09336
modelloabbonamentod_26	150.10311	104.76734	1.43	0.1534	0.14666
modellomistod_26	20.17128	97.53369	0.21	0.8364	0.02181
grafica_31	-7.55275	13.60613	-0.56	0.5794	-0.03327
ambientazione_31	6.67559	17.79394	0.38	0.7079	0.0249
storia_31	-14.97552	11.79175	-1.27	0.2055	-0.07841
pvp_31	-20.98746	10.04882	-2.09	0.0379	-0.11889
interazione_31	18.97545	15.56262	1.22	0.2241	0.07023
perfomance_35	-25.7292	12.98381	-1.98	0.0488	-0.10794
online_38	3.90077	1.03274	3.78	0.0002	0.24121
soddisfazione_42	19.86908	12.60424	1.58	0.1164	0.09014

Una volta indivuato il set di regressori di partenza si è proceduto alla creazione del modello composto da 22 regressori. L'indice R-quadro mostra un valore di 0,45 (ricordando che si partiva da uno 0,5 con tutte le variabili possibili): il passaggio da più di 40 variabili a 22 ha portato a una perdita di bontà del modello pari al 5%.

Analizzando però il test t, che indica la probabilità che il coefficiente di regressione sia 0, vediamo come alcune variabili hanno un'alta probabilità di non influire nel modello.

Vengono quindi eliminati tutti i regressori con valore p-value superiore a 0.2, ottenendo così 10 regressori.

Da 22 a 10 Regressori

Stime dei parametri. R-quadro=0.43 R-quadro corr=0.4					
Variabile	Stima dei parametri	Errore standard	Valore t	Pr > |t|	Stima standardizzata
Interc.	-6411.90716	3452.0448	-1.86	0.0645	0
Eta_2	7.38742	4.483	1.65	0.1008	0.0856
PCm_10	64.5338	34.4732	1.87	0.0625	2.30685
Consolem_10	62.92301	34.31081	1.83	0.068	2.26068
ngiochi_11	-29.40233	10.87371	-2.7	0.0074	-0.14298
durata_21	0.60502	0.07617	7.94	<.0001	0.42665
modelloabbonamentod_26	117.26992	55.51808	2.11	0.0358	0.11458
pvp_31	-24.67134	9.25939	-2.66	0.0083	-0.13976
perfomance_35	-28.24615	12.39233	-2.28	0.0236	-0.1185
online_38	4.70297	0.90171	5.22	<.0001	0.29081
soddisfazione_42	15.09803	11.9302	1.27	0.207	0.0685

Passando da 22 a 10 regressori si ha una ulteriore perdita della bontà del modello pari al 2%.

Troviamo però ancora un regressore, soddisfazione, con una probabilità del 20% di essere ininfluente nel modello, poiché potrebbe avere un coefficiente tendente a 0. Infatti il valore di stima standarizzata mostra come all'aumentare o diminuire della soddisfazione, la durata aumenta o diminuisca solamente del 6%.

Si procede quindi all'eliminazione di questo regressore, e all'analisi della multicollinearità tra i diversi regressori rimasti.

Analisi di Multicollinearità

Coefficienti di correlazione di Pearson

Prob > |r| rispetto a H0: Rho=0

Numero di osservazioni

	Eta_2	PCm_10	Consolem_10	ngiochi_11	durata_21	modelloabbonamentod_26	pvp_31	perfomance_35	online_38
Eta_2	1	-0.01246	0.01306	0.02089	0.10778	0.17962	-0.03262	0.08436	0.13958
Eta_2		0.8458	0.8385	0.7444	0.0971	0.0047	0.6107	0.1872	0.0286
	246	246	246	246	238	246	246	246	246
PCm_10	-0.01246	1	-0.99915	-0.12518	0.12192	-0.03605	-0.05462	-0.09461	0.09161
PCm_10	0.8458		<.0001	0.0499	0.0604	0.5736	0.3937	0.139	0.152
	246	246	246	246	238	246	246	246	246
Consolem_10	0.01306	-0.99915	1	0.126	-0.12226	0.04062	0.05585	0.09252	-0.09167
Consolem_10	0.8385	<.0001		0.0484	0.0597	0.526	0.3831	0.1479	0.1517
	246	246	246	246	238	246	246	246	246
ngiochi_11	0.02089	-0.12518	0.126	1	0.10117	0.0428	-0.02955	0.11801	-0.08551
ngiochi_11	0.7444	0.0499	0.0484		0.1196	0.504	0.6447	0.0646	0.1813
	246	246	246	246	238	246	246	246	246
durata_21	0.10778	0.12192	-0.12226	0.10117	1	-0.12506	-0.13223	-0.03708	0.16617
durata_21	0.0971	0.0604	0.0597	0.1196		0.054	0.0415	0.5692	0.0102
	238	238	238	238	238	238	238	238	238
modelloabbonamentod_26	0.17962	-0.03605	0.04062	0.0428	-0.12506	1	0.15062	0.1246	0.22888
modelloabbonamentod_26	0.0047	0.5736	0.526	0.504	0.054		0.0181	0.0509	0.0003
	246	246	246	246	238	246	246	246	246
pvp_31	-0.03262	-0.05462	0.05585	-0.02955	-0.13223	0.15062	1	0.04758	0.20044
pvp_31	0.6107	0.3937	0.3831	0.6447	0.0415	0.0181		0.4575	0.0016
	246	246	246	246	238	246	246	246	246
perfomance_35	0.08436	-0.09461	0.09252	0.11801	-0.03708	0.1246	0.04758	1	0.0757
perfomance_35	0.1872	0.139	0.1479	0.0646	0.5692	0.0509	0.4575		0.2368
	246	246	246	246	238	246	246	246	246
online_38	0.13958	0.09161	-0.09167	-0.08551	0.16617	0.22888	0.20044	0.0757	1
online_38	0.0286	0.152	0.1517	0.1813	0.0102	0.0003	0.0016	0.2368	
	246	246	246	246	238	246	246	246	246

Troviamo solo un valore che indica una forte correlazione tra 2 variabili: PCm e Consolem con un valore tendente a 1. Viene così eliminata una variabile.

Scelta finale del modello

Stime dei parametri. R-quadro= 0,41 R-quadro corr= 0,39					
Variabile	Stima dei parametri	Errore standard	Valore t	Pr > |t|	Stima standardizzata
Interc.	-5.27357	204.01701	-0.03	0.9794	0
Eta_2	7.40909	4.51387	1.64	0.1021	0.08585
PCm_10	1.1443	1.45366	0.79	0.432	0.0409
ngiochi_11	-26.11574	10.7072	-2.44	0.0155	-0.127
durata_21	0.61461	0.0764	8.05	<.0001	0.43342
modelloabbonamentod_26	130.33547	55.50781	2.35	0.0197	0.12734
pvp_31	-24.39149	9.32019	-2.62	0.0095	-0.13817
perfomance_35	-27.89031	12.36329	-2.26	0.025	-0.11701
online_38	4.88867	0.88825	5.5	<.0001	0.3023

Stime dei parametri. R-quadro= 0.41 R-quadro corr= 0.39					
Variabile	Stima dei parametri	Errore standard	Valore t	Pr > |t|	Stima standardizzata
Interc.	110.52219	141.24971	0.78	0.4347	0
Eta_2	7.2894	4.50758	1.62	0.1072	0.08447
ngiochi_11	-27.13965	10.61911	-2.56	0.0112	-0.13197
durata_21	0.62104	0.0759	8.18	<.0001	0.43795
modelloabbonamentod_26	129.7926	55.45762	2.34	0.0201	0.12681
pvp_31	-24.79818	9.29817	-2.67	0.0082	-0.14047
perfomance_35	-28.63067	12.31727	-2.32	0.021	-0.12012
online_38	4.94875	0.88423	5.6	<.0001	0.30601

Dopo aver eliminato i regressori soddisfazione e consolem il livello di bontà del modello scende di un ulteriore 2% e si ferma a 0,41. Una variabile risulta ancora ininfulente, con un valore del test t superiore a 0,4. Si procede quindi alla eliminazione della variabile pcm e si ottiene un modello a 7 regressori, con bontà quasi invariata.

Viene quindi effettuato un controllo finale sui sette regressori tramite l'analisi della correlazione tra variabili.

Multicollinearità finale

Coefficienti di correlazione di Pearson Prob > |r| rispetto a H0: Rho=0 Numero di osservazioni							
	1	ngiochi_11	durata_21	modelloabbonamentod_26	pvp_31	perfomance_35	online_38
Eta_2		0.02089	0.10778	0.17962	-0.03262	0.08436	0.13958
Eta_2		0.7444	0.0971	0.0047	0.6107	0.1872	0.0286
	246	246	238	246	246	246	246
ngiochi_11	0.02089	1	0.10117	0.0428	-0.02955	0.11801	-0.08551
ngiochi_11	0.7444		0.1196	0.504	0.6447	0.0646	0.1813
	246	246	238	246	246	246	246
durata_21	0.10778	0.10117	1	-0.12506	-0.13223	-0.03708	0.16617
durata_21	0.0971	0.1196		0.054	0.0415	0.5692	0.0102
	238	238	238	238	238	238	238
modelloabbonamentod_26	0.17962	0.0428	-0.12506	1	0.15062	0.1246	0.22888
modelloabbonamentod_26	0.0047	0.504	0.054		0.0181	0.0509	0.0003
	246	246	238	246	246	246	246
pvp_31	-0.03262	-0.02955	-0.13223	0.15062	1	0.04758	0.20044
pvp_31	0.6107	0.6447	0.0415	0.0181		0.4575	0.0016
	246	246	238	246	246	246	246
perfomance_35	0.08436	0.11801	-0.03708	0.1246	0.04758	1	0.0757
perfomance_35	0.1872	0.0646	0.5692	0.0509	0.4575		0.2368
	246	246	238	246	246	246	246
online_38	0.13958	-0.08551	0.16617	0.22888	0.20044	0.0757	1
online_38	0.0286	0.1813	0.0102	0.0003	0.0016	0.2368	
	246	246	238	246	246	246	246

Dall'analisi della multicollinearità tramite coefficiente di correlazione risulta che le variabili non sono tra loro correlate.

Il modello a 7 regressori è un modello accettabile, anche se con un livello di bontà scarso pari a 0,41

Selection Stepwise

Stime dei parametri. R-quadro= 0.41 R-quadro corr= 0.39					
Variabile	Stima dei parametri	Errore standard	Valore t	Pr > |t|	Stima standardizzata
Interc.	110.52219	141.24971	0.78	0.4347	0
Eta_2	7.2894	4.50758	1.62	0.1072	0.08447
ngiochi_11	-27.13965	10.61911	-2.56	0.0112	-0.13197
durata_21	0.62104	0.0759	8.18	<.0001	0.43795
modelloabbonamentod_26	129.7926	55.45762	2.34	0.0201	0.12681
pvp_31	-24.79818	9.29817	-2.67	0.0082	-0.14047
perfomance_35	-28.63067	12.31727	-2.32	0.021	-0.12012
online_38	4.94875	0.88423	5.6	<.0001	0.30601

La scelta dei regressori attraverso la selezione stepwise ci porta ad avere 7 regressori.

Questi regressori sono i medesimi trovati mediante l'analisi manuale, per cui non verrà fatto nessun controllo su di essi.

Interpretazione del modello

Stime dei parametri. R-quadro= 0.41 R-quadro corr= 0.39					
Variabile	Stima dei parametri	Errore standard	Valore t	Pr > |t|	Stima standardizzata
Interc.	110.52219	141.24971	0.78	0.4347	0
Eta_2	7.2894	4.50758	1.62	0.1072	0.08447
ngiochi_11	-27.13965	10.61911	-2.56	0.0112	-0.13197
durata_21	0.62104	0.0759	8.18	<.0001	0.43795
modelloabbonamentod_26	129.7926	55.45762	2.34	0.0201	0.12681
pvp_31	-24.79818	9.29817	-2.67	0.0082	-0.14047
perfomance_35	-28.63067	12.31727	-2.32	0.021	-0.12012
online_38	4.94875	0.88423	5.6	<.0001	0.30601

L'interpretazione del modello avviene mediante l'analisi della stima standardizzata per ogni regressore.

Troviamo alcuni regressori che influiscono positivamente sulla variabile dipendente e alcuni che influiscono negativamente.

Il regressore che spiega maggiormente la variabile è, come prevedibile, il tempo che un giocatore dedica a un altro videogioco. Al variare di una unità della variabile, la variabile varia con una percentuale pari al 43%: questo significa che i giocatori che restano maggiormente su un gioco è probabile che lo faranno anche su un altro. Un ulteriore regressore che influenza positivamente con una buona percentuale la variabile dipendente è il tempo che un giocatore dedica all'online rispetto al singleplayer: in questo caso la variazione è pari al 30%. Coloro che giocano maggiormente in multiplayer saranno portati a restare su un gioco maggiormente. Vi sono altri due regressori che spiegano una variazione positiva, seppur in maniera inferiore, e sono l'età, con un 8% e il modello di business ad abbonamento per un 12%: questo significa che un pubblico giovane tende ad avere una rotazione di giochi maggiori; nel caso del modello ad abbonamento, invece, coloro che utilizzano giochi con questo modello sono portati a rimanere sul gioco maggiormente.

Vi sono infine 3 regressori che influiscono negativamente sulla variabile dipendente con una percentuale circa del 13%: numero di giochi, pvp e performance. Secondo la prima variabile maggiormente una persona gioca a molteplici videogames minore sarà la durata sul singolo gioco, la seconda invece indica che maggiormente è apprezzato il sistema di player vs player minor tempo si dedicherà a un gioco, poiché i giocatori cercheranno sempre nuove sfide innovative per misurarsi con altri giocatori, infine la terza variabile indica che coloro che posseggono un PC ad alto potenziale saranno portati a giocare meno su un singolo gioco, probabilmente per testare sempre la potenza del proprio PC sui nuovi giochi.

Conclusione

E’ stato individuato tramite analisi dei singoli fattori e selezione stepwise un modello a 7 regressori tra loro non fortemente correlati che spiegano circa il 41% del fenomeno della durata di un giocatore in un videogioco. Il valore è sicuramente basso, poiché un buon modello dovrebbe spiegare almeno il 60% del fenomeno, ma i dati in nostro possesso non erano sufficienti.

Non è stato affrontato il metodo dell’analisi fattoriale per superare problemi relativi alla multicollinearità, poiché poche variabili erano collegate tra loro.

Regressione Logistica

Il modello di regressione logistica ci permette di studiare un comportamento dicotomico di un individuo spiegato da una variabile bernoulliana sulla base di altre variabili dette regressori.

Nel nostro caso si analizzerà un servizio accessorio collegato alla possibilità che i giocatori scarichino i videogiochi direttamente dal sito. Questo servizio è importante in quanto può essere parte fondamentale del fatturato del servizio di gaming online.

Coefficienti di correlazione di Pearson Prob > |r| rispetto a H0: Rho=0 Numero di osservazioni	
	scaricare_41
Eta_2 Eta_2	0.09348 0.1438 246
free_ore_giorno_4 free_ore_giorno_4	-0.10918 0.0875 246
ngiochi_11 ngiochi_11	-0.12651 0.0475 246
notte_12 notte_12	0.09718 0.1285 246
cartacredito_28 cartacredito_28	0.08903 0.1639 246
compravendita_29 compravendita_29	0.1858 0.0034 246
interazione_31 interazione_31	-0.12327 0.0535 246
perfomance_35 perfomance_35	-0.11394 0.0745 246
conoscenzapc_36 conoscenzapc_36	0.09957 0.1193 246
online_38 online_38	0.15094 0.0178 246
single_38 single_38	-0.15248 0.0167 246
fps_39 fps_39	-0.10347 0.1106 239

Scelta dei regressori

Come nel modello di regressione lineare i regressori sono stati scelti in base alla correlazione con la variabile dipendente.

In questo caso è stato individuato un set di regressori di partenza formato da 12 variabili. La scelta è stata effettuata su una correlazione con coefficiente tendente a 0,1 o maggiore.

Modello a 12 regressori

Test dell'ipotesi globale nulla: BETA=0				Associazione di probabilità previste e risposte osservate			
Test	**Chi-quadrato**	**DF**	**Pr > ChiQuadr**	**Percentuale concordante**	78.5	**Di di Somers**	0.57
Rapp. verosim.	51.0463	12	<.0001	**Percentuale discordante**	21.2	**Gamma**	0.57
Score	43.5976	12	<.0001	**Percentuale legato**	0.2	**Tau-a**	0.27
Wald	34.962	12	0.0005	**Coppie**	13158	**c**	0.79

Analisi delle stime di massima verosimiglianza					
Parametro	**Stima**	**Errore standard**	**Wald Chi-quadrato**	**Pr > ChiQuadr**	**Stima standardizzata**
Intercept	88.6532	1054.3	0.0071	0.933	
Eta_2	0.0296	0.035	0.7184	0.3967	0.077
free_ore_giorno_4	-0.1227	0.0626	3.8383	0.0501	-0.175
ngiochi_11	-0.1689	0.0739	5.2211	0.0223	-0.1962
notte_12	0.0146	0.0096	2.3	0.1294	0.1476
cartacredito_28	0.1002	0.0488	4.2139	0.0401	0.1835
compravendita_29	0.9497	0.3708	6.5592	0.0104	0.2388
interazione_31	-0.355	0.1202	8.7208	0.0031	-0.3149
perfomance_35	-0.2191	0.0976	5.0416	0.0247	-0.2181
conoscenzapc_36	0.2568	0.1015	6.403	0.0114	0.2378
online_38	-0.8471	10.543	0.0065	0.936	-12.4508
single_38	-0.8648	10.543	0.0067	0.9346	-12.6775
fps_39	-0.0435	0.0274	2.5168	0.1126	-0.146

Il modello mostra un valore di bontà espresso dal test chi quadro buono. Infatti la probabilità che i regressori e la variabile dipendente siano indipendenti è bassissima.

Un ulteriore indicatore sul modello è la percentuale di concordant rispetto alla percentuale di discordant, dati sia dai valori % che dagli indicatori Di di Somers, Gamma, Tau-a e c.

Si nota però che il test t chi quadro mostra valori elevati per alcuni regressori.

Si procede quindi all'analisi di multicollinearità per analizzare la correlazione tra le singole variabili.

Esiste infatti una forte correlazione tra online e single (come prevedibile poiché % di una stessa domanda).

Si procede all'esclusione di online_38 e del regressore eta_2 poiché mostra un valore del test t elevato (0,4).

Da 10 a 8 Regressori

Analisi delle stime di massima verosimiglianza						
Parametro	DF	Stima	Errore standard	Wald Chi-quadrato	Pr > ChiQuadr	Stima standardizzata
Intercept	1	4.5657	1.4361	10.1077	0.0015	
free_ore_giorno_4	1	-0.1349	0.0613	4.8496	0.0277	-0.1925
ngiochi_11	1	-0.1705	0.0733	5.408	0.02	-0.1981
notte_12	1	0.0175	0.00927	3.5481	0.0596	0.1767
cartacredito_28	1	0.104	0.0484	4.6125	0.0317	0.1905
compravendita_29	1	0.9322	0.3709	6.3179	0.012	0.2344
interazione_31	1	-0.3553	0.1208	8.6582	0.0033	-0.3152
perfomance_35	1	-0.2151	0.0972	4.8954	0.0269	-0.2142
conoscenzapc_36	1	0.2699	0.1003	7.2412	0.0071	0.2499
single_38	1	-0.0177	0.00621	8.0763	0.0045	-0.2589
fps_39	1	-0.0458	0.0273	2.8165	0.0933	-0.1536

Analisi delle stime di massima verosimiglianza						
Parametro	DF	Stima	Errore standard	Wald Chi-quadrato	Pr > ChiQuadr	Stima standardizzata
Intercept	1	3.363	1.283	6.8703	0.0088	
free_ore_giorno_4	1	-0.1025	0.0577	3.1632	0.0753	-0.1449
ngiochi_11	1	-0.1311	0.0686	3.6496	0.0561	-0.1528
cartacredito_28	1	0.0693	0.045	2.3783	0.123	0.1281
compravendita_29	1	0.9876	0.353	7.8259	0.0052	0.2472
interazione_31	1	-0.2708	0.1071	6.3996	0.0114	-0.2404
perfomance_35	1	-0.2211	0.0947	5.4549	0.0195	-0.2192
conoscenzapc_36	1	0.2309	0.0928	6.1955	0.0128	0.212
single_38	1	-0.0142	0.00582	5.9518	0.0147	-0.2096

Dopo aver eliminato le due variabili sono state effettuate 4 diverse prove, a 10 regressori, a 8, a 7 e a 5 eliminando ogni volta valori del test chi quadro superiori a 5.

Da 7 a 5 Regressori

Analisi delle stime di massima verosimiglianza						
Parametro	DF	Stima	Errore standard	Wald Chi-quadrato	Pr > ChiQuadr	Stima standardizzata
Intercept	1	3.529	1.2713	7.7062	0.0055	
free_ore_giorno_4	1	-0.102	0.0574	3.1549	0.0757	-0.1442
ngiochi_11	1	-0.1211	0.068	3.1734	0.0748	-0.1412
compravendita_29	1	0.9441	0.3482	7.3528	0.0067	0.2363
interazione_31	1	-0.2736	0.1065	6.5977	0.0102	-0.2428
perfomance_35	1	-0.2015	0.0928	4.7139	0.0299	-0.1997
conoscenzapc_36	1	0.2356	0.0921	6.5427	0.0105	0.2163
single_38	1	-0.015	0.00576	6.7971	0.0091	-0.2219

Analisi delle stime di massima verosimiglianza						
Parametro	DF	Stima	Errore standard	Wald Chi-quadrato	Pr > ChiQuadr	Stima standardizzata
Intercept	1	2.5686	1.1682	4.8349	0.0279	
compravendita_29	1	0.8465	0.3386	6.2499	0.0124	0.2119
interazione_31	1	-0.2491	0.1031	5.8339	0.0157	-0.221
perfomance_35	1	-0.2149	0.0907	5.6099	0.0179	-0.213
conoscenzapc_36	1	0.2318	0.0913	6.4508	0.0111	0.2129
single_38	1	-0.0165	0.00565	8.5352	0.0035	-0.2438

Le diverse prove hanno portato infine a un modello a 5 regressori con valori inferiori a 0,3.

E' stato poi effettutato un test di bontà generale sui quattro diversi modelli.

Bontà del modello

Test dell'ipotesi globale nulla: BETA=0				Associazione di probabilità previste e risposte osservate			
Test	**Chi-quadrato**	**DF**	**Pr > ChiQuadr**	**Percentuale concordante**	78	**Di di Somers**	0.56
Rapp. verosim.	49.6935	10	<.0001	**Percentuale discordante**	21.8	**Gamma**	0.56
Score	42.3737	10	<.0001	**Percentuale legato**	0.2	**Tau-a**	0.26
Wald	34.66	10	0.0001	**Coppie**	13158	**c**	0.78
Test dell'ipotesi globale nulla: BETA=0				**Associazione di probabilità previste e risposte osservate**			
Test	**Chi-quadrato**	**DF**	**Pr > ChiQuadr**	**Percentuale concordante**	74.3	**Di di Somers**	0.49
Rapp. verosim.	38.8197	8	<.0001	**Percentuale discordante**	25.4	**Gamma**	0.49
Score	34.9857	8	<.0001	**Percentuale legato**	0.3	**Tau-a**	0.23
Wald	29.7729	8	0.0002	**Coppie**	14040	**c**	0.75
Test dell'ipotesi globale nulla: BETA=0				**Associazione di probabilità previste e risposte osservate**			
Test	**Chi-quadrato**	**DF**	**Pr > ChiQuadr**	**Percentuale concordante**	74.1	**Di di Somers**	0.49
Rapp. verosim.	36.4014	7	<.0001	**Percentuale discordante**	25.6	**Gamma**	0.49
Score	33.2928	7	<.0001	**Percentuale legato**	0.3	**Tau-a**	0.23
Wald	28.6326	7	0.0002	**Coppie**	14040	**c**	0.74
Test dell'ipotesi globale nulla: BETA=0				**Associazione di probabilità previste e risposte osservate**			
Test	**Chi-quadrato**	**DF**	**Pr > ChiQuadr**	**Percentuale concordante**	71.9	**Di di Somers**	0.44
Rapp. verosim.	29.8014	5	<.0001	**Percentuale discordante**	27.8	**Gamma**	0.44
Score	27.6642	5	<.0001	**Percentuale legato**	0.3	**Tau-a**	0.21
Wald	24.5255	5	0.0002	**Coppie**	14040	**c**	0.72

Il test di bontà mostra come si abbia una perdita di percentuali concordant da 78% a 72%. Indicatori come c e Tau-a diminuiscono di pochi punti, mentre Gamma e Di di Somers perdono anche 10 punti percentuali.

L'indicatore p-value sul test chi-quadro rimane praticamente invariato.

Scelta del modello

Analisi delle stime di massima verosimiglianza						
Parametro	DF	Stima	Errore standard	Wald Chi-quadrato	Pr > ChiQuadr	Stima standardizzata
Intercept	1	4.5657	1.4361	10.1077	0.0015	
free_ore_giorno_4	1	-0.1349	0.0613	4.8496	0.0277	-0.1925
ngiochi_11	1	-0.1705	0.0733	5.408	0.02	-0.1981
notte_12	1	0.0175	0.00927	3.5481	0.0596	0.1767
cartacredito_28	1	0.104	0.0484	4.6125	0.0317	0.1905
compravendita_29	1	0.9322	0.3709	6.3179	0.012	0.2344
interazione_31	1	-0.3553	0.1208	8.6582	0.0033	-0.3152
perfomance_35	1	-0.2151	0.0972	4.8954	0.0269	-0.2142
conoscenzapc_36	1	0.2699	0.1003	7.2412	0.0071	0.2499
single_38	1	-0.0177	0.00621	8.0763	0.0045	-0.2589
fps_39	1	-0.0458	0.0273	2.8165	0.0933	-0.1536

Test dell'ipotesi globale nulla: BETA=0				Associazione di probabilità previste e risposte osservate			
Test	Chi-quadrato	DF	Pr > ChiQuadr	Percentuale concordante	78	Di di Somers	0.56
Rapp. verosim.	49.6935	10	<.0001	Percentuale discordante	21.8	Gamma	0.56
Score	42.3737	10	<.0001	Percentuale legato	0.2	Tau-a	0.26
Wald	34.66	10	0.0001	Coppie	13158	c	0.78

Si è scelto infine un modello a 10 regressori, poiché presenta da un lato alti valori di concordant e degli indici di bontà.

Inoltre osservando la stima standardizzata anche quei regressori con p-value del test chi-quadro superiori a 0,3 spiegano sempre il fenomeno con una percentuale superiore del 10%.

Prendendo per esempio il regressore FPS si noti come una variazione della variabile comporti una spiegazione pari al 15% del fenomeno nel senso opposto.

Stepwise

Analisi delle stime di massima verosimiglianza						
Parametro	DF	Stima	Errore standard	Wald Chi-quadrato	Pr > ChiQuadr	Stima standardizzata
Intercept	1	2.2638	1.2511	3.274	0.0704	
Eta_2	1	0.0657	0.0328	4.0223	0.0449	0.1728
ngiochi_11	1	-0.1505	0.0699	4.633	0.0314	-0.1746
compravendita_29	1	0.861	0.3491	6.0845	0.0136	0.2155
interazione_31	1	-0.3074	0.1119	7.5424	0.006	-0.2735
single_38	1	-0.0138	0.00597	5.3495	0.0207	-0.1996

Test dell'ipotesi globale nulla: BETA=0				Associazione di probabilità previste e risposte osservate			
Test	Chi-quadrato	DF	Pr > ChiQuadr	Percentuale concordante	72.2	Di di Somers	0.447
Rapp. verosim.	27.1351	5	<.0001	Percentuale discordante	27.5	Gamma	0.448
Score	25.0381	5	0.0001	Percentuale legato	0.3	Tau-a	0.207
Wald	22.6592	5	0.0004	Coppie	12118	c	0.723

Per effettuare un confronto è stata utilizzata anche la procedura di selezione stepwise, per cui sono stati individuati 5 regressori.

Rispetto alla soluzione a 10 regressori abbiamo una minor percentuale di concordant e valori di bontà, seppur di poco, inferiori a quelli del modello scelto.

Rimane comunque una buona scelta anche il modello individuato dalla selezione stepwise.

Interpretazione del modello

Analisi delle stime di massima verosimiglianza						
Parametro	DF	Stima	Errore standard	Wald Chi-quadrato	Pr > ChiQuadr	Stima standardizzata
Intercept	1	4.5657	1.4361	10.1077	0.0015	
free_ore_giorno_4	1	-0.1349	0.0613	4.8496	0.0277	-0.1925
ngiochi_11	1	-0.1705	0.0733	5.408	0.02	-0.1981
notte_12	1	0.0175	0.00927	3.5481	0.0596	0.1767
cartacredito_28	1	0.104	0.0484	4.6125	0.0317	0.1905
compravendita_29	1	0.9322	0.3709	6.3179	0.012	0.2344
interazione_31	1	-0.3553	0.1208	8.6582	0.0033	-0.3152
perfomance_35	1	-0.2151	0.0972	4.8954	0.0269	-0.2142
conoscenzapc_36	1	0.2699	0.1003	7.2412	0.0071	0.2499
single_38	1	-0.0177	0.00621	8.0763	0.0045	-0.2589
fps_39	1	-0.0458	0.0273	2.8165	0.0933	-0.1536

L'interpretazione del modello avviene mediante l'analisi della stima standardizzata per ogni regressore.

Troviamo alcuni regressori che influiscono positivamente sulla variabile dipendente e alcuni che influiscono negativamente.

I regressori che influiscono positivamente sono 4.

La variabile notte indica che maggiormente un giocatore dedica il proprio tempo al gaming on-line durante questo periodo del giorno, maggiore sarà la probabilità che acquisti i giochi dallo shop: questo aumento sarà pari circa al 18%. La variabile carta di credito indica invece una variazione sulla probabilità che si verifichi il fenomeno di acquisto del 19%: maggiore è l'utilizzo di questa modalità di pagamento per i videogiochi on-line maggiore sarà la probabilità di acquisto. Altri due regressori che influiscono positivamente sono la compravendita di oggetti e la conoscenza del pc, con variazioni sulla probabilità del 23/24%: una maggior propensione all'acquisto deriva, per la prima variabile, a seguito di una propensione allo scambio di oggetti on-line tramite utilizzo di denaro reale, per la seconda invece, a seguito di una miglior conoscenza del pc, e quindi alla praticità di installare i giochi senza cd.

Variabili che influiscono negativamente sono invece fps, con una variazione del 15%. che indica una minor propensione da parte di coloro che apprezzano maggiormente il genere; ore libere al giorno, con una variazione del 19% indica che coloro che hanno maggior tempo libero preferiscono acquistare su internet tramite spedizione o in negozio, numero di giochi, con una variazione del 19% mostra che maggiore è il numero di giochi on-line praticati minore sarà la propensione ad acquistare i giochi on-line; singleplayer, con una variazione del 25% indica che chi spende maggior tempo per i giochi on-line preferisce acquistarli in negozio o su internet; performance, con una variazione del 21% indica che chi ha un pc a maggior potenziale preferisce maggiormente i giochi in scatola; infine interazione, con una variazione del 31% indica che coloro che preferiscono giochi in cui vi sia forte legame con altri giocatori sono meno propensi a scaricare i giochi, poiché si tratta solitamente di browser game.

Conclusioni

Sono stati individuati due modelli, a 10 regressori e a 5 regressori, che spiegano in maniera soddisfacente il fenomeno di possibilità di acquisto dei videogames direttamente on-line.

Bisogna però dire che, come nella regressione lineare, anche in questo modello le variabili non spiegano completamente il fenomeno, ma solamente in parte.

Il modello può dare solo indicazioni di massima.

Conclusioni finali

- Sono state individuate le caratteristiche fondamentali del target di riferimento, quali età, professione, tempo libero, tempo dedicato al gaming, momento della giornata in cui si concentra il maggior numero di giocatori, al fine di conoscere quale possibile clientela avrà il servizio.
- Sono state individuate le caratteristiche del core business, indicando quale piattaforma di gioco viene maggiormente utilizzata, quali siano i genere da proporre e con quale modalità di business, incluso il prezzo da praticare laddove vi sia richiesto.
- Sono stati individuati inoltre quali servizi accessori possono essere offerti, come chatvocali, forum e possibilità di scaricare i giochi dal sito internet.
- Sono state individuate quali modalità di pagamento adottare maggiormente: carte prepagate e carta di credito.
- E' stato costruito un modello di analisi fattoriale al fine di spiegare il comportamento dei giocatori attraverso solo 9 variabili (factor loadings)
- E' stato costruito un modello di regressione lineare per spiegare una variabile fondamentale del servizio quale la durata dell'impegno di un giocatore su in videogioco
- Infine è stato costruito un modello di regressione logistica per spiegare il comportamento dei giocatori rispetto alla possibilità di scaricare i videogiochi direttamente dal sito internet, servizio che può portare un buon margine di guadagno.

Punti forti dell’analisi

- Buona conoscenza del target, tramite analisi univariata, bivariata e fattoriale, al fine di poter intraprendere le giuste politiche di marketing.
- Ottima conoscenza delle caratteristiche dei generi offribili per raggruppare il maggior numero di giocatori possibili.
- Buona conoscenza dei servizi accessori per aumentare la soddisfazione dei clienti.

Punti di debolezza

- L'analisi spiega bene la base di partenza da cui creare il servizio, ma non la gestione a causa di modelli incompleti.
- Modello di regressione lineare mostra un valore di varianza spiegata pari al 41%.
- Modello di regressione logistica risulta leggermente complicato a causa dei notevoli contrasti nel segno delle variabili.

Limiti riscontrati

- Le variabili sono maggiormente indicative per dare informazioni di massima
- Il questionario è stato costruito in un momento in cui si conosceva l'analisi univariata e bivariata, ma non l'analisi fattoriale e i modelli di regressione. Per questo si ritiene a posteriori che il questionario non sia perfettamente idoneo al lavoro svolto
- Numero di survey limitato a 246.
- Inapplicabilità di modelli di campionamento possibili a causa del numero limitato di survey: si poteva effettuare un campionamento in base al numero di survey ricevute e il numero di utenti iscritti sul forum
- Utilizzo di un canale indiretto e impersonale per la raccolta dei dati quali l'utilizzo di questionario informatico.

Soluzioni

- Creazione di un nuovo questionario mirato al fine di creare nuovi modelli di regressione per capire meglio il comportamento dei giocatori.
- Creazione del servizio e utilizzo dei dati interni per effettuare analisi sul target diretto: oggettivizzazione dei dati.
- Ampliare il numero di intervistati.