

GameStarter

Andrea Mattei [546139]

Luna Noemi Puccia [550968]

Fabio Carannante [546779]

Abstract

Gamestarter nasce dalla curiosità di esplorare il mondo del crowdfunding, un modello di finanziamento che permette di realizzare un proprio progetto grazie al sostegno volontario di investitori privati. La piattaforma di crowdfunding più frequentata è Kickstarter.com il quale adotta una filosofia all-or-nothing: i fondi vengono trasferiti al progettista solo qualora venga effettivamente raggiunto l'obiettivo prefissato. Vista l'enormità di Kickstarter (più di 415.000 progetti), abbiamo limitato la nostra analisi a una sola categoria, Games, la più importante su Kickstarter (885 milioni di dollari incassati), e ad un solo anno, il 2018. Le metodologie che abbiamo applicato si prestano comunque ad un eventuale ampliamento dell'analisi.

Introduzione

Il sito intende offrire uno squarcio sul fenomeno ormai diffusissimo del crowdfunding che anno dopo anno sta conquistando sempre più il mondo della imprenditoria. Abbiamo scelto come piattaforma di riferimento, in quanto la più usata, Kickstarter.com, un sito che permette un tipo di finanziamento reward-based, in cui ogni investitore può scegliere tra varie opzioni di contributo e ricevere in cambio la ricompensa associata, e in particolare sulla sua categoria più prolifica: i Games. Abbiamo scelto quest'ultima per due ragioni principali: la prima è perché, come è stato detto prima è la categoria che funziona di più tra tutte; la seconda è perché intendiamo far avvicinare le generazioni di oggi a questo mondo ancora troppo sconosciuto in Italia.

Per aiutarci a tale scopo abbiamo creato dei grafici che rappresentano al meglio ciò che è il crowdfunding e come si sia diffuso oramai in tutto il mondo. Non essendoci dati disponibili in internet, abbiamo pensato fosse il caso di estrarli grazie all'aiuto di Scrapy in modo tale che i dati potessero essere i più aggiornati possibili.

Stato dell'arte

Essendo un fenomeno in rapida espansione, il crowdfunding viene scelto come argomento da molti per la pubblicazione di articoli e creazione di siti web. Un aspetto sul quale abbiamo però voluto distinguerci è l'aver scelto una singola categoria e non trattarlo nella sua complessità.

<https://www.crowd-funding.cloud/it/i-neri-del-crowdfunding-italiano-nel-2017-874.asp>

su questo sito ad esempio sono riportati i dati riguardanti l'andamento delle varie tipologie di crowdfunding e non si scende nel particolare delle singole categorie di progetti. Inoltre l'argomento è trattato a livello nazionale (Italia) e non a livello internazionale come abbiamo invece scelto noi.

Per quanto riguarda i siti che trattano i games nel crowdfunding abbiamo trovato diversi esempi come: <https://www.gamesindustry.biz/articles/2018-01-11-video-games-crowdfunding-raised-usd25m-in-2017>

Oppure

nei quali però i dati sono stati presi da report annuali. Noi grazie alla nostra personale estrazione di dati da Kickstarter.com siamo riusciti ad avere i dati mensili più recenti e a dare uno sguardo all'anno ancora in corso.

Modello dei Dati

• Estrazione dei dati

I dati che abbiamo analizzato sono stati ottenuti tramite dei web crawler scritti in Python, un framework open source che permette di programmare i propri spider personalizzati. Siamo partiti da Kicktraq.com (<http://www.kicktraq.com/>), un sito che contiene svariate informazioni su tutti i progetti di Kickstarter, attivi o terminati. L'appoggio su Kicktraq è stato fondamentale in quanto Kickstarter cancella periodicamente le pagine di vecchi progetti falliti dal suo motore di ricerca interno, rendendo difficile automatizzare un crawler. Per ognuna delle sottocategorie della categoria Games abbiamo lanciato il nostro primo web crawler per ottenere informazioni come il nome del progetto, la quantità di denaro raccolta, lo stato di residenza del progettista, gli url di Kickstarter alla pagina del progetto e alla bio dell'autore, eccetera. Una volta scaricati i dati dei progetti di tutto il 2018, abbiamo lanciato un secondo crawler su tutti i link alle bio degli autori che abbiamo ottenuto nel primo JSON, in modo da ottenere un secondo file con alcuni dati sui progettisti, ad esempio il loro numero di amici su Facebook nel caso avessero collegato il loro account, oppure il numero di progetti altrui che hanno sostenuto. Dopodiché abbiamo convertito i due file JSON in CSV tramite un semplice tool online (<http://www.convertcsv.com/json-to-csv.htm>). Per completezza e trasparenza abbiamo incluso i file Python dei nostri due spider nella cartella Web_Crawling del progetto.

• Manipolazione dei dati

La maggior parte delle modifiche ai dati sono state fatte direttamente nei file Python dei web crawler, in modo tale da ottenere dati già pronti per essere caricati direttamente dall'output. Le modifiche principali che abbiamo attuato sono stati cambi di tipo (per rendere i numeri che erano stati scrapati come stringhe effettivamente dei numeri) e manipolazioni di stringhe, per eliminare dati indesiderati e mantenere solo informazioni rilevanti. Abbiamo poi aggiunto una parte di codice che fungesse da convertitore di valuta, in modo da convertire tutti gli importi in dollari americani per poterli poi confrontare. Una volta convertiti i dati in formato CSV, su Excel abbiamo semplicemente eliminato alcuni record relativi a progetti precedenti il 2018 e aggiunto un campo ID con numeri in sequenza.

• Schema del Database

Il nostro database arriva dunque a essere composto da due grandi tabelle, una relativa ai progetti e una relativa agli autori. A sinistra abbiamo la tabella "progetti", a destra in alto la tabella "autori". Le due tabelle sono collegate tra loro dai campi "author_link" in "progetti" e "author_url" in "autori", che contiene lo stesso valore, ossia il link alla bio del creatore di quel progetto. Nel database abbiamo anche una terza, piccola tabella (in basso a destra), che contiene alcuni dati presi da un articolo del sito Polygon.com, che abbiamo utilizzato nell'introduzione del nostro progetto.

#	Nome	Tipo
1	ID	int(4)
2	name	varchar(79)
3	urlkick	varchar(109)
4	subcategory	varchar(15)
5	locationcity	varchar(28)
6	locationstate	varchar(14)
7	author	varchar(40)
8	currency	varchar(3)
9	pledged_amount	int(8)
10	pledged_converted	int(7)
11	status	varchar(10)
12	trending_to	int(8)
13	trending_to_converted	int(7)
14	author_link	varchar(120)
15	description	varchar(150)
16	backers_number	int(5)
17	start_date	varchar(52)
18	end_date	varchar(52)
19	year	int(4)
20	update_number	int(2)
21	macrocategory	varchar(5)
22	goal	int(7)
23	pledge_percent	int(6)
24	estimated_percent	int(6)

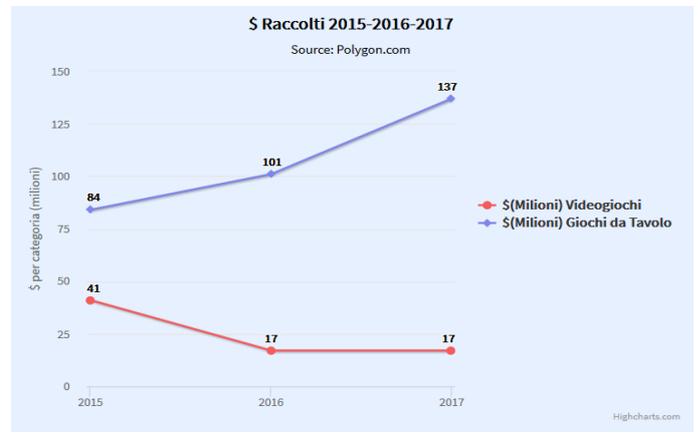
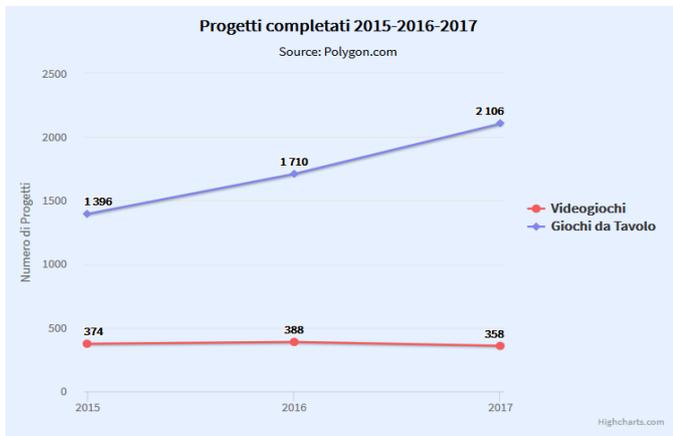
#	Nome	Tipo
1	ID	int(4)
2	nome_autore	varchar(40)
3	author_url	varchar(120)
4	true_name	varchar(59)
5	fb_friends	int(4)
6	fb_conn	varchar(3)
7	link_created	varchar(64)
8	link_backed	varchar(63)
9	projects_created	int(2)
10	projects_backed	int(3)

#	Nome	Tipo
1	ID	int(1)
2	Anno	int(4)
3	Successful_VG	int(3)
4	Successful_Tabletop	int(4)
5	Milions_VG	varchar(5)
6	Milions_Tabletop	varchar(6)

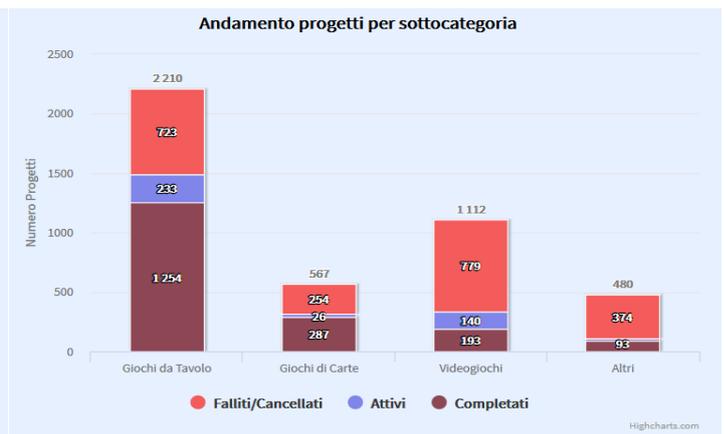
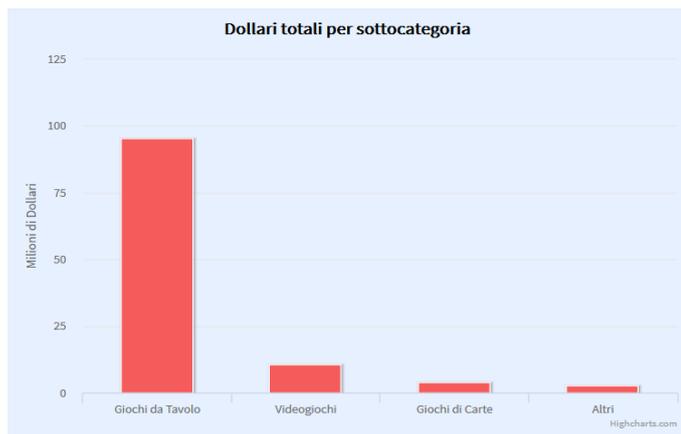
Analisi dei Dati

I grafici che abbiamo inserito nel sito sono stati realizzati grazie a Highcharts e Geocharts.

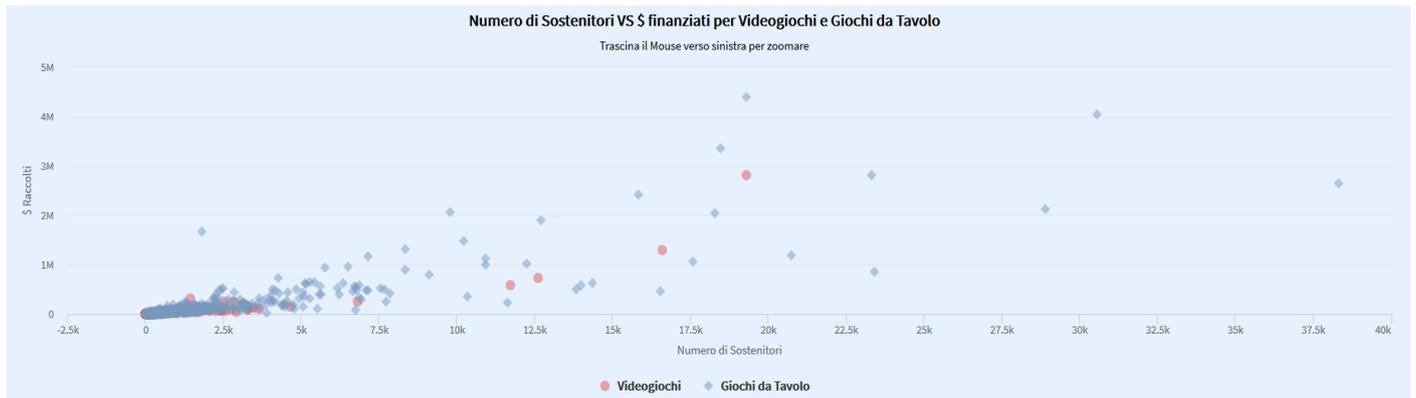
Con la prima coppia di grafici abbiamo voluto analizzare sin da subito la macrocategoria dei games, scindendola nelle sue due sottocategorie più prolifiche: i videogiochi e i giochi da tavolo. In particolare abbiamo voluto dimostrare quanti progetti nel corso degli anni 2015, 2016, 2017 siano stati effettivamente completati e quanti soldi siano stati investiti grazie al finanziamento crowdfunding, concludendo che c'è stato un netto predominio dei giochi da tavolo sui videogiochi.



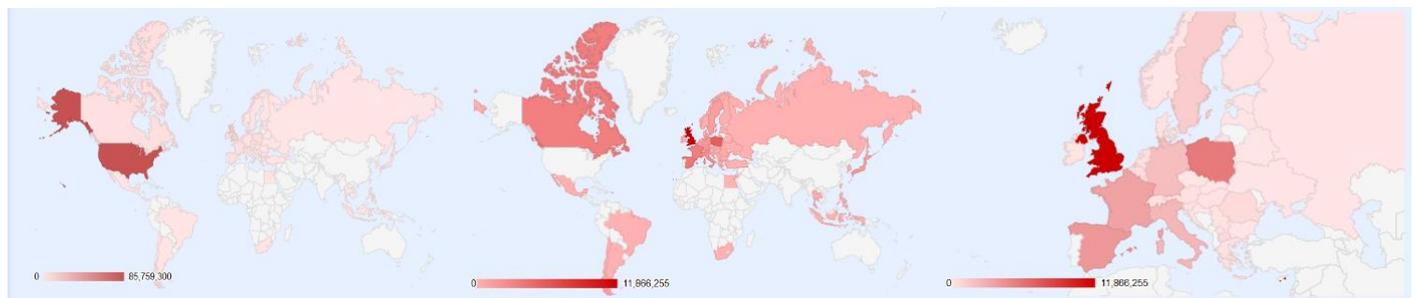
Con la coppia successiva invece intendiamo volgere lo sguardo al presente e far vedere come le varie sottocategorie dei games si stiano comportando nel corso dell'anno corrente, notando il solito dominio dei videogiochi e dei giochi da tavolo su tutte le altre e anche che il trend sembri rimanere molto simile a quello dell'anno precedente.



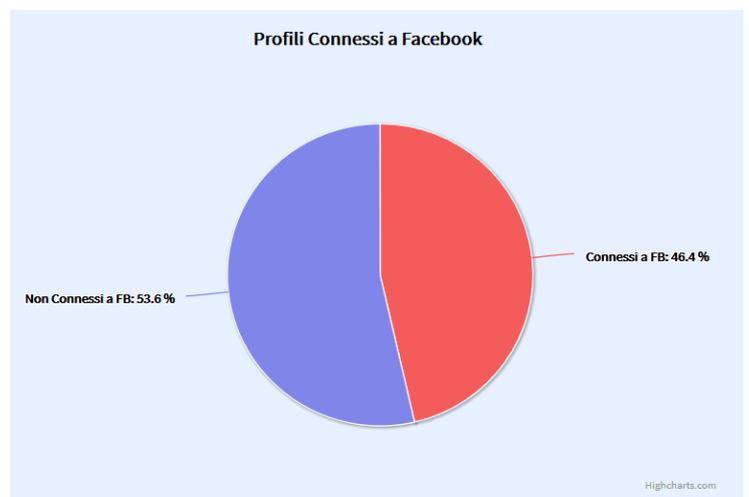
Il successivo grafico di tipo scatter plot raccoglie tutti i progetti di videogiochi e giochi da tavolo mettendo in relazione il numero di investitori che hanno sostenuto il progetto con la somma totale raccolta. Vogliamo dunque dimostrare come non sia vero che più sostenitori il progetto riceve e più sono i guadagni. Oltre a ciò abbiamo voluto portare qualche esempio dei progetti singoli più riusciti.



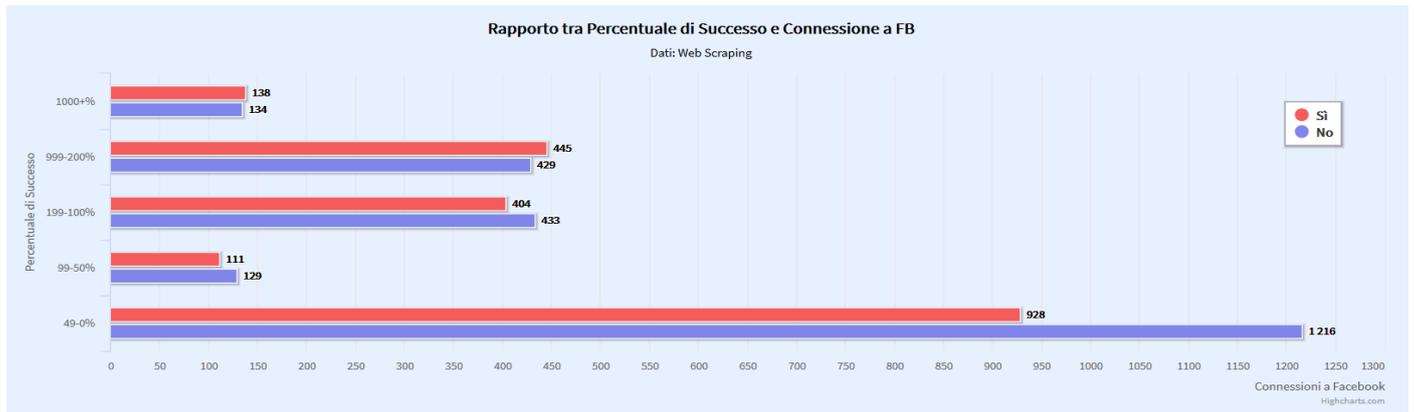
Abbiamo scelto poi di inserire 3 grafici da Geocharts nei quali mettiamo da parte le sottocategorie per dare uno sguardo agli attuali dati della categoria 'games'. Volevamo vedere in quale stato venissero prodotti più giochi via crowdfunding. Avendo potuto osservare come gli Stati Uniti surclassino tutti i rimanenti stati abbiamo deciso di inserire altre due cartine che potessero spiegare al meglio la situazione nel resto del mondo: nel primo si ha una visione mondiale senza però comprendere gli Stati Uniti; nel secondo invece viene riprodotta esclusivamente la situazione europea.



Abbiamo poi visto come il crowdfunding possa essere strettamente collegato ai social network. Per questo abbiamo realizzato dapprima un semplice grafico a torta in cui è possibile vedere quanti dei profili degli utenti sono collegati all'account Facebook; successivamente abbiamo voluto stabilire se la connessione social rappresenti davvero un vantaggio per gli utenti di Kickstarter.



Nel primo grafico è rappresentato il rapporto tra percentuale di successo e numero di progetti i cui autori hanno l'account connesso a Facebook. Si può vedere come nelle fasce la distribuzione sia abbastanza regolare, al contrario, nella fascia più bassa ci sia un netto scarto tra profili NON connessi e profili connessi, e inoltre come le due fasce più alte siano le uniche in cui i profili connessi siano in numero maggiore, seppur di poco.

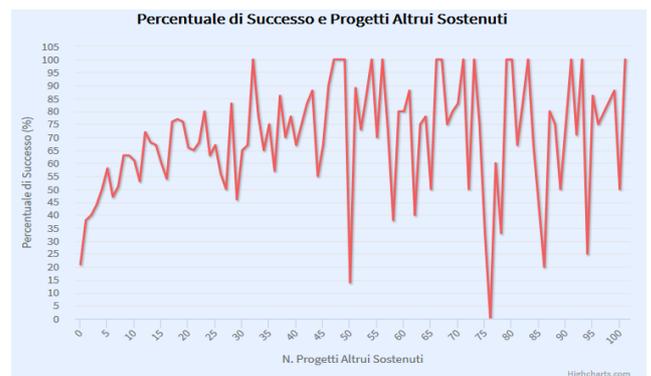


Nel secondo grafico abbiamo voluto dimostrare se il numero di amici su Facebook dell'utente conta o no nella buona riuscita del progetto. Abbiamo dunque messo in relazione le fasce di successo con il numero medio di amici su Facebook, notando come gli autori dei progetti che hanno avuto successo abbiano un numero medio di amici Facebook nettamente superiore rispetto a quelli che ne hanno uno minore.

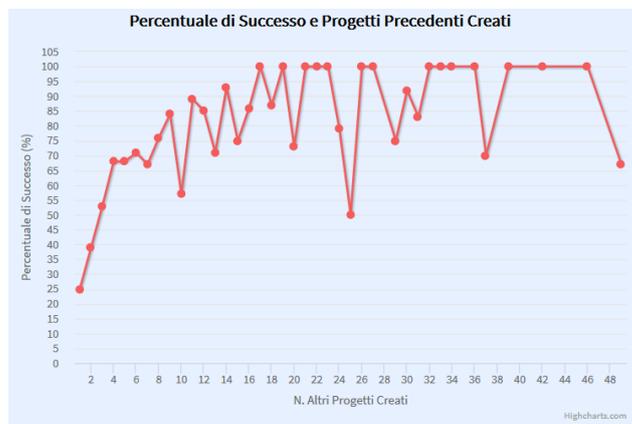


Per concludere, abbiamo preso in considerazione i progettisti, per vedere quanto possano incidere i progetti altrui sostenuti e i progetti creati dallo stesso progettista in precedenza.

Il primo grafico ci mostra la percentuale di successo per progetti altrui: possiamo notare che sostenendo più progetti la probabilità che il proprio progetto vada a buon fine sia maggiore, probabilmente perché un investitore si sente più sicuro ad investire verso un progettista che a sua volta ha investito e così via.



Il secondo grafico ci mostra la probabilità di successo per il numero di altri progetti creati in precedenza: i progettisti “alle prime armi” hanno meno probabilità di riuscire ad arrivare ai fondi necessari per completare la propria campagna. Più si va avanti con il numero di progetti creati, più la probabilità aumenta: grazie all’esperienza, un progettista riesce a capire quale sia l’idea migliore per convincere un investitore a collaborare al proprio progetto e come pubblicizzare la propria campagna.



Conclusioni e Possibili Sviluppi

Con Gamestarter abbiamo provato a dare una panoramica sul fenomeno del crowdfunding in un particolare settore che ritenevamo potesse interessare in modo particolare ai giovani, per i quali questo modello di business è più che mai rilevante, vista la carenza di lavoro stipendiato degli ultimi anni.

Riteniamo che un progetto di questo tipo si presti molto bene a sviluppi futuri: si potrebbe estendere l’analisi altre categorie di Kickstarter, oppure estendere il dataset temporalmente arrivando a progetti fino al 2009, o ancora effettuare un’analisi geografica più specifica, osservando non soltanto a livello di nazioni, ma anche a livello di città. Queste operazioni potrebbero essere effettuate con delle semplici modifiche ai nostri spider. Con l’aiuto di un economista, si potrebbero poi formulare teorie più accurate delle nostre sulle leggi che regolano il crowdfunding, probabilmente arrivando a poter prevedere la percentuale di successo di un nuovo progetto con discreta precisione.