

# Discount Information

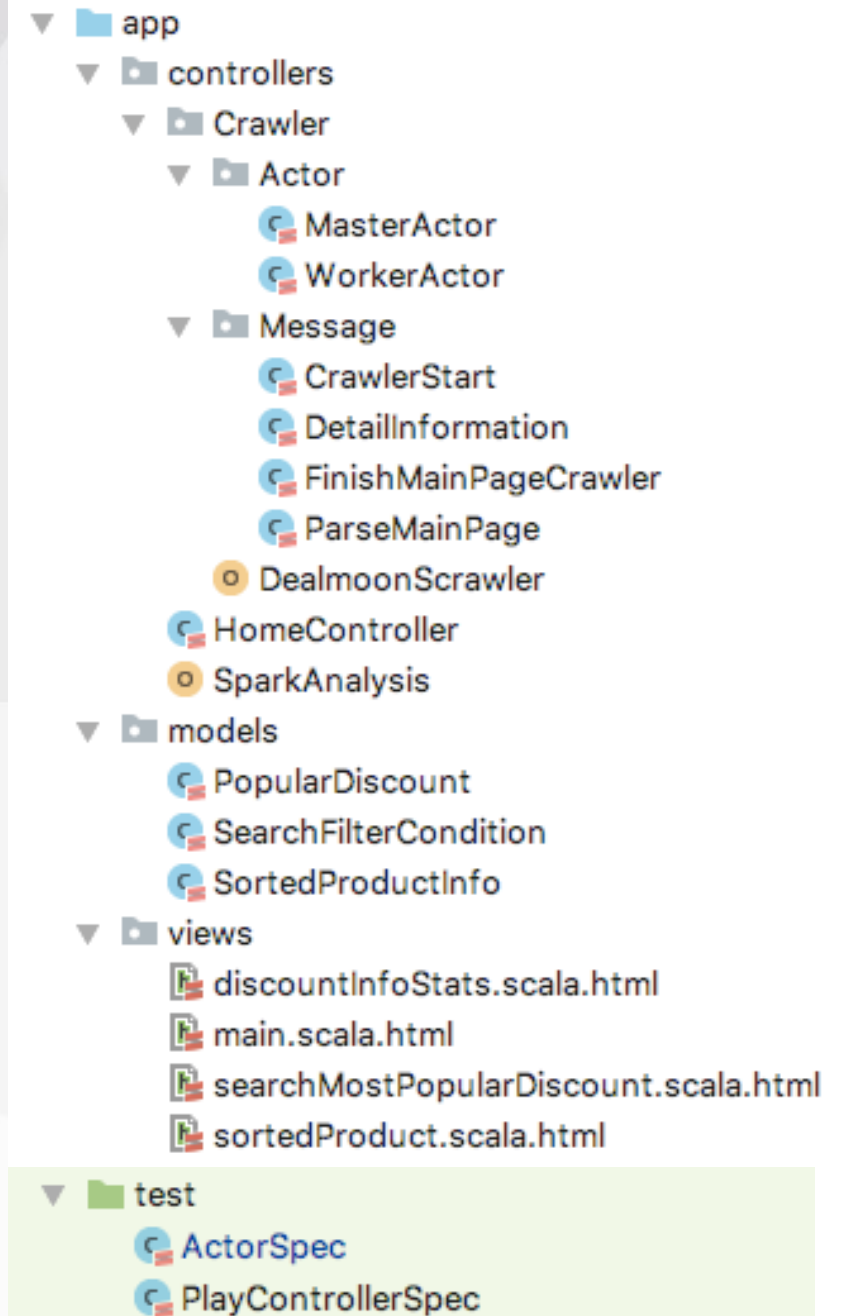
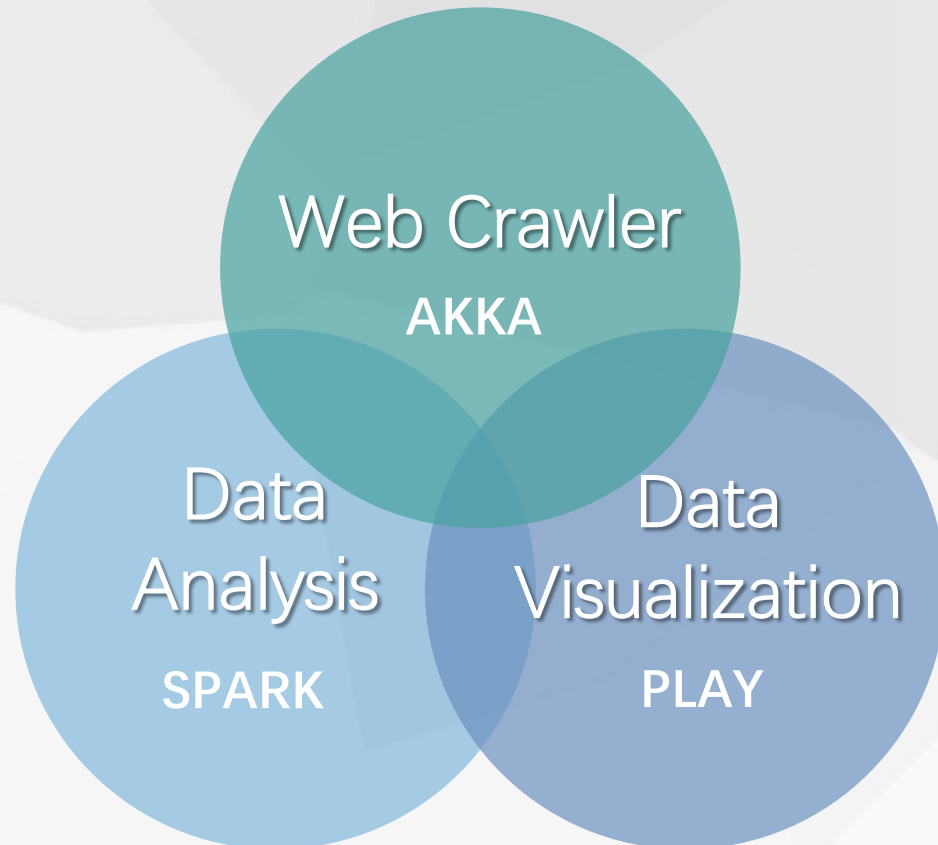
Searching and Sharing Platform



# PART 1

## **Design Methodology & Code Structure**

## Project Composition




# Web Page Structure Introduction

## Deal Information Item In Index Page

Containing deal description,  
detail page link, platform,  
favorites, shares, comments

01

Exclusive




12 mins. ago

**Enjoy 20% off + GWP with Advanced Genifique Sensitive Serum purchase @ Lancôme**

- Dealmoon Exclusive!
- Lancôme offers [enjoy 20% off with \\$75+ Advanced Genifique Sensitive Serum purchase](#) via coupon code "DMGREENMONDAY".
- Plus, receive 4 samples & bag on \$75 orders.
- Plus, receive a full-size Hydra Zen Day Cream on \$125 orders, a **\$48 value**.
- Free 2-day shipping on any order.
- Deal ends 12/12.

9 5 Lancôme Beauty

HOT



12 mins. ago

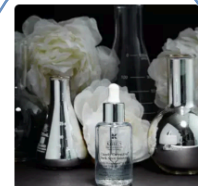
**Get \$20 off with \$65+ all Facial Serums Purchase @ Kiehl's**

- Kiehl's offers [Get \\$20 off with \\$65+ all Facial Serums purchase](#) via coupon code "JINGLE".
- And receive a free gift box with \$85+ purchase.
- Choose 3 samples at check out.
- Free shipping on orders over \$50.
- Deal ends 12/16

## Discounted Product Item In Detailed Page

Containing discounted  
product name, brand, original  
price, discounted price

02



**\$50.00**  
Kiehl's Clearly Corrective™ Dark...



**\$47.00**  
Kiehl's Midnight Recovery Concentrate



**\$62.00**  
Kiehl's Powerful-Strength Line-...

Show More ▾

## Categories :

"Clothing-Jewelry-Bags" -> 80  
"Beauty" -> 80  
"Electronics-Computers" -> 88  
"At-Home" -> 94  
"Baby-Kids" -> 66  
"Sports-Outdoors" -> 19

## Data Dimension :

Detail Page URL  
Discount Description  
Platform  
Product Title  
Discounted Price  
Original Price  
Category  
Favorites  
Shares  
Comments

## Actor and Task Structure

### Task 1: Create Child Worker

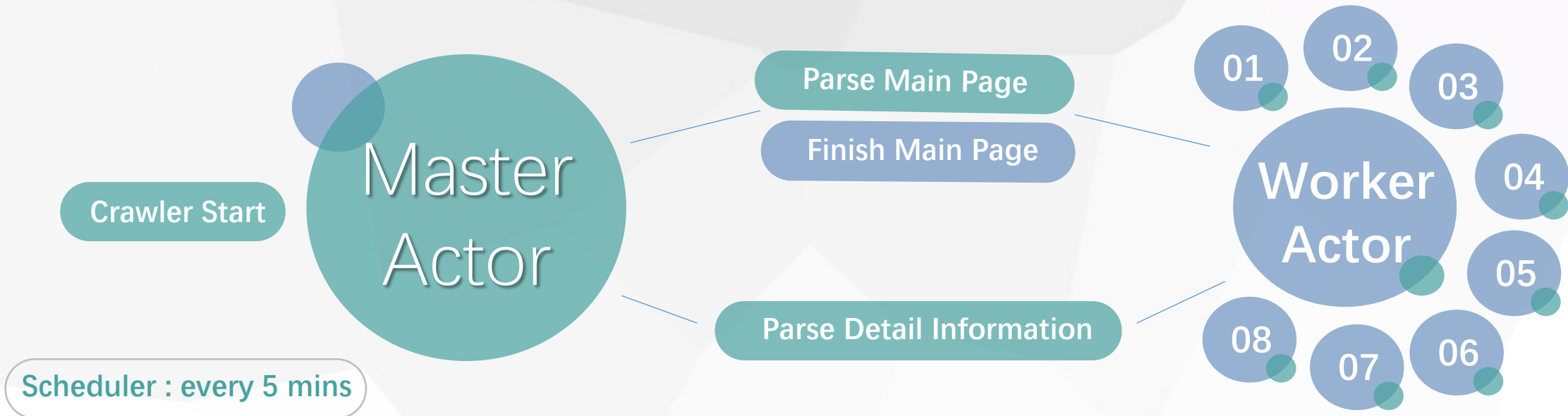
Create and assign one child worker for each page of six categories, the total number of worker is  $\text{Sum}(\text{pages})$

### Task 2: Index Page Crawler

Get and save the detailed information page URLs and deal description by crawling the index page

### Task 3: Detailed Deal Information Crawler

After finishing the main page crawler task, worker will be assigned to crawler detailed discounted product information



# Data Structure & Spark Processing

Data Scale : 50101

URL	Discount	Platform	Title	DiscountedPrice	OriginalPrice	Category	Favorites	Shares	Comments
https://www.dealmoon...	Up to 50% Off + Free...	Reebok	Reebok CrossFit Regional 2...	\$14.97	\$28.00	Sports-Outdoors	3	<null>	<null>
https://www.dealmoon...	Up to 50% Off + Free...	Reebok	Reebok UFC Fight Gear Figh...	\$19.99	\$25.00	Sports-Outdoors	3	<null>	<null>
https://www.dealmoon...	Up to 50% Off + Free...	Reebok	Reebok 2018 CrossFit Games...	\$24.97	\$38.00	Sports-Outdoors	3	<null>	<null>
https://www.dealmoon...	Up to 50% Off + Free...	Reebok	Reebok LES MILLS™ Dual Ble...	\$29.97	\$45.00	Sports-Outdoors	3	<null>	<null>
https://www.dealmoon...	Up to 50% Off + Free...	Reebok	Reebok LOW CUT SOCKS – 3 P...	\$6.97	\$14.00	Sports-Outdoors	3	<null>	<null>
https://www.dealmoon...	Up to 50% Off + Free...	Reebok	Reebok Club MEMT Wide 4E	\$29.97	\$65.00	Sports-Outdoors	3	<null>	<null>
https://www.dealmoon...	Up to 50% Off + Free...	Reebok	Reebok Classic Vector Tee	\$16.97	\$25.00	Sports-Outdoors	3	<null>	<null>
https://www.dealmoon...	Up to 50% Off + Free...	Reebok	Reebok Classic Leather	\$74.97	\$85.00	Sports-Outdoors	3	<null>	<null>
https://www.dealmoon...	Up to 50% Off + Free...	Reebok	Reebok Rehbok T-Shirt	\$24.97	\$35.00	Sports-Outdoors	3	<null>	<null>
https://www.dealmoon...	Up to 50% Off + Free...	Reebok	Reebok Classic Slide	\$19.97	\$25.00	Sports-Outdoors	3	<null>	<null>
https://www.dealmoon...	Up to 50% Off + Free...	Reebok	Reebok AT SUPER 3.0 STEALTH	\$109.99	\$130.00	Sports-Outdoors	3	<null>	<null>
https://www.dealmoon...	Up to 50% Off + Free...	Reebok	Reebok CrossFit Regionals ...	\$12.97	\$42.00	Sports-Outdoors	3	<null>	<null>
https://www.dealmoon...	Up to 50% Off + Free...	Reebok	Reebok Running Tight	\$29.97	\$39.99	Sports-Outdoors	3	<null>	<null>
https://www.dealmoon...	Up to 50% Off + Free...	Reebok	Reebok Revenge Plus Archive	\$104.97	\$115.00	Sports-Outdoors	3	<null>	<null>
https://www.dealmoon...	From \$5.99	Urban Outfi...	Champion & UO Colorblock L...	\$55.00	<null>	Sports-Outdoors	83	<null>	4
https://www.dealmoon...	From \$5.99	Urban Outfi...	Champion Reverse Weave Bas...	\$19.99	\$30.00	Sports-Outdoors	83	<null>	4
https://www.dealmoon...	From \$5.99	Urban Outfi...	Urban Outfitters Champion ...	\$25.00	\$35.00	Sports-Outdoors	83	<null>	4
https://www.dealmoon...	From \$5.99	Urban Outfi...	Urban Outfitters Champion ...	\$55.00	<null>	Sports-Outdoors	83	<null>	4
https://www.dealmoon...	From \$5.99	Urban Outfi...	Urban Outfitters Champion ...	\$40.00	<null>	Sports-Outdoors	83	<null>	4
https://www.dealmoon...	From \$5.99	Urban Outfi...	Champion Reverse Weave ALL...	\$29.00	\$55.00	Sports-Outdoors	83	<null>	4

CSV File

Operations :

Filter  
Sort  
Order  
Select  
Sum  
Map  
Reduce  
Count  
Distinct  
Collect  
Take  
Limit

Spark Data frame

URL	Discount	Platform	Picture	Title	DiscountedPrice	OriginalPrice	Category	Favorites	Shares	Comments
https://www.dealm...	Up to 80% Off	Sierra Trading Post	Marmot Camp EZ Ja...	\$49.99	null	Sports-Outdoors	22	null	null	null
https://www.dealm...	40% Off + Free Sh...	Reebok	Reebok Classic Le...	\$51.00	\$85.00	Sports-Outdoors	29	null	1	null
https://www.dealm...	Up to 25% Off + F...	Eastbay	Nike Rally Loose ...	\$45.00	\$60.00	Sports-Outdoors	13	null	1	null
https://www.dealm...	Up to 50% Off + F...	Columbia Sportswear	Columbia Men's St...	\$29.99	\$34.99	Sports-Outdoors	64	null	1	null
https://www.dealm...	25% Off Over \$100	Under Armour	Under Armour Wome...	\$33.99	\$45.00	Sports-Outdoors	72	null	17	null
https://www.dealm...	From \$4.8	Lord & Taylor	Puma Classic Logo...	\$10.50	\$25.00	Sports-Outdoors	22	null	null	null
https://www.dealm...	Daily Update	Walmart	Nintendo Labo Rob...	\$59.99	\$79.99	Sports-Outdoors	6113	null	1604	null
https://www.dealm...	Up to 60% Off	Moosejaw	Moosejaw Women's ...	\$186.75	\$249.00	Sports-Outdoors	7	null	null	null
https://www.dealm...	From \$9+Free Ship...	Lululemon	Lululemon Princip...	\$99.00	\$128.00	Sports-Outdoors	119	null	16	null
https://www.dealm...	Extra 30% Off	6PM.com	The North Face Nu...	\$52.15	\$149.00	Sports-Outdoors	37	null	13	null
https://www.dealm...	Up to 50% Off	FinishLine.com	Nike Men's Nike S...	\$59.99	\$140.00	Sports-Outdoors	180	null	12	null
https://www.dealm...	From \$1199.99	Amazon.com	Apple MacBook Pro...	\$1199.99	\$1499.99	Electronics-Compu...	6	null	5	null
https://www.dealm...	null	World Wide Stereo	JBL Everest 110 W...	\$49.95	\$129.95	Electronics-Compu...	31	null	7	null
https://www.dealm...	A7ii for \$998	Amazon.com	Sony Alpha a6300 ...	\$748.00	null	Electronics-Compu...	102	null	14	null
https://www.dealm...	\$399.99	Amazon.com	Garmin Fenix 5S -...	\$399.99	\$549.99	Electronics-Compu...	5	null	4	null
https://www.dealm...	Extra 20% Off + 3...	Reebonz	Prada Spazzolato ...	\$1055.20	\$1990.00	Clothing-Jewelry-...	18	null	5	null
https://www.dealm...	Up to 60% Off	SSENSE	MCQ ALEXANDER MCQ...	\$119.00	\$195.00	Clothing-Jewelry-...	94	null	6	null
https://www.dealm...	27% Off	H&M	H&M Off-the-shoul...	\$23.35	\$39.99	Clothing-Jewelry-...	340	null	71	null
https://www.dealm...	Up to 60% Off	Gilt	Roger Vivier Embe...	\$699.99	\$1900.00	Clothing-Jewelry-...	4	null	null	null
https://www.dealm...	Up To 60% Off	Barneys New York	Chloe Roy Small L...	\$1000.30	\$949.00	Clothing-Jewelry-...	661	null	60	null



## PART 2

# User Stories & Unit Test

## User Story 1

### Search Most Popular Deal

Select Category and Matrix

Get Top 10 Deals Description and Platform sorted by matrix

category

Beauty



matrix

By Favorites



Search

Discount

Platform

Up to 50% off

Silk'n

Up to 78% off

Sephora.com

Free 12-pc Gift

Nordstrom

Up to 70% Off + Extra 20% Off

Perfumania

New customers receive 10% off

Charlotte Tilbury

Up to 83% off

Kiko Milano

Dealmoon Exclusive deluxe samples

Nordstrom

25% Off

macys.com

New Arrivals!

Saks Fifth Avenue

Buy 2 get 1 free + Free Shipping

Nordstrom



## User Story 2

### Search Discounted Product

Select Keyword & Category

**keyword**

**category**

Clothing-Jewelry-Bags ⬆⬇⬆

Search

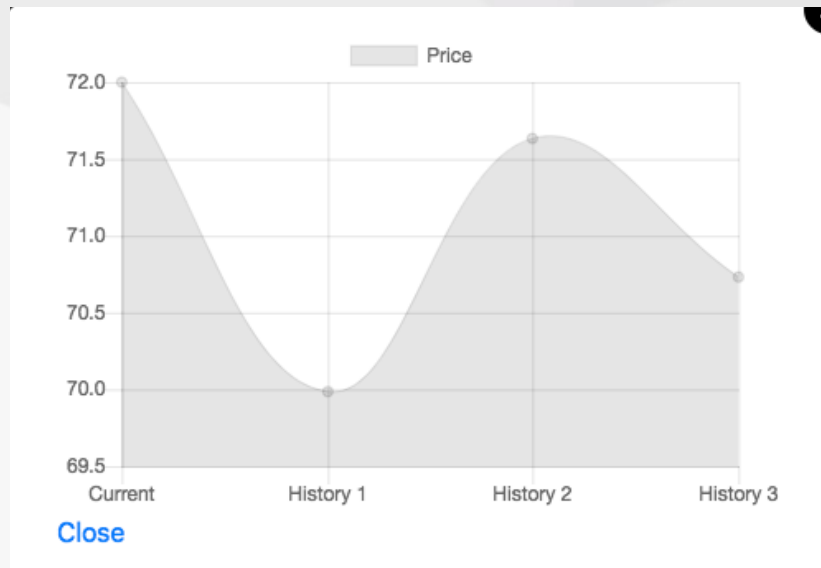
Get All Discounted Products Contains Keyword

Title	Discounted Price	Original Price	Platform	Operation	Operation
Golden Goose Deluxe Brand Superstar distressed leather sneakers	£270.83	£325.00	NET-A-PORTER UK	Add to Favourite	Trend Graph
Golden Goose Deluxe Brand Superstar distressed leather and suede sneakers	£258.33	£310.00	NET-A-PORTER UK	Add to Favourite	Trend Graph
Golden Goose Deluxe Brand Superstar glittered distressed leather sneakers	£258.33	£310.00	NET-A-PORTER UK	Add to Favourite	Trend Graph
Golden Goose Deluxe Brand Superstar distressed metallic leather and suede sneakers	£250.00	£300.00	NET-A-PORTER UK	Add to Favourite	Trend Graph
Jimmy Choo Diamond metallic-leather sneakers	\$895.00	null	Selfridges	Add to Favourite	Trend Graph
Balenciaga Triple S sneakers	\$890.00	null	Mytheresa	Add to Favourite	Trend Graph
ASOS Vans X Disney SK8-Hi magic sneakers at asos.com	\$80.00	null	ASOS	Add to Favourite	Trend Graph

## User Story 3

### Get Trend Graph and Follow of Discounted Product

View the Trend Graph of Price



Add to Wish List and send the email

Wish List :

email

Send Email

Name

Discount Information

Valentino Garavani Rockstud metallic sneakers

Price Have Been Discounted To \$657.00

#### An Update Of Your Recent Favourite Items



xingyao wu <sparkprojecttest@outlook.com>  
to wu.x@husky.neu.edu, me ▾

3:48 PM (0 minutes ago)

#### Recent Update Of Your Item

##### Discount Information:

Item	Discount Information
Valentino The Rockstud cotton-jersey sweatshirt	\$870.00
ALEXACHUNG Printed cotton-jersey T-shirt	£37.50
Jimmy Choo Diamond metallic-leather sneakers	\$895.00

Reply

Reply all

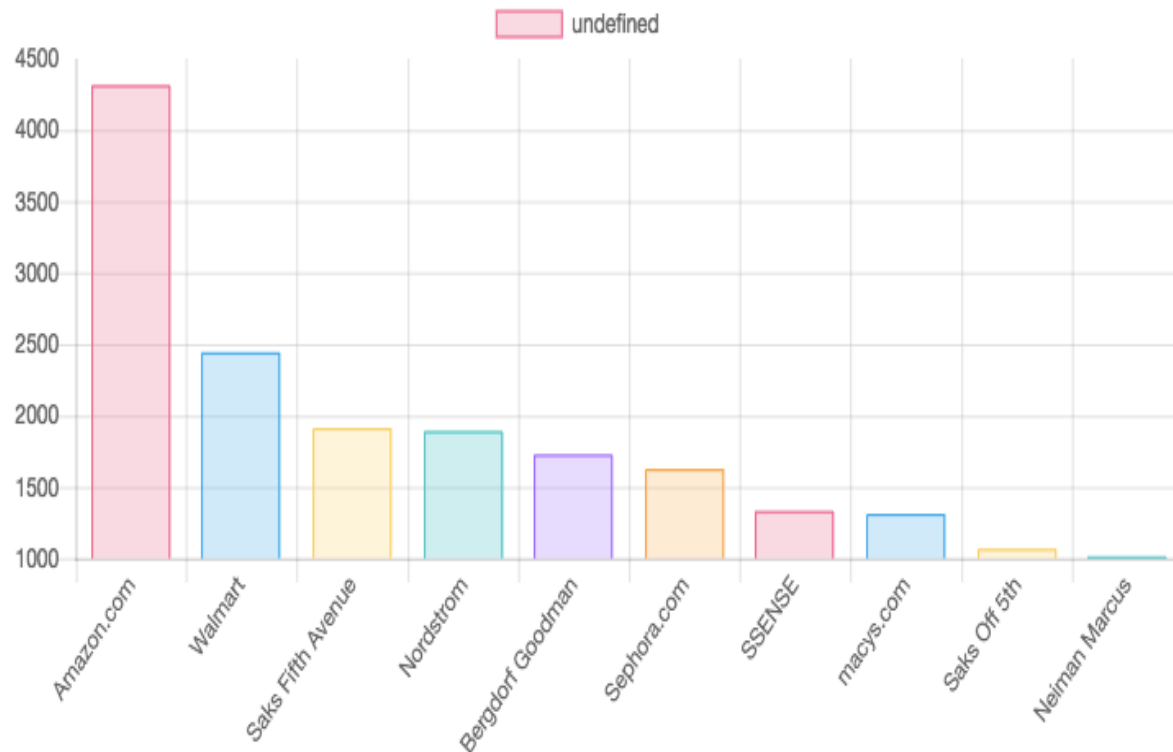
Forward

## User Story 4

### Get Statistics Information on Most Published and Popular Platforms

#### Top 10 Most Published Platforms

Top 10 Most Post Platforms:



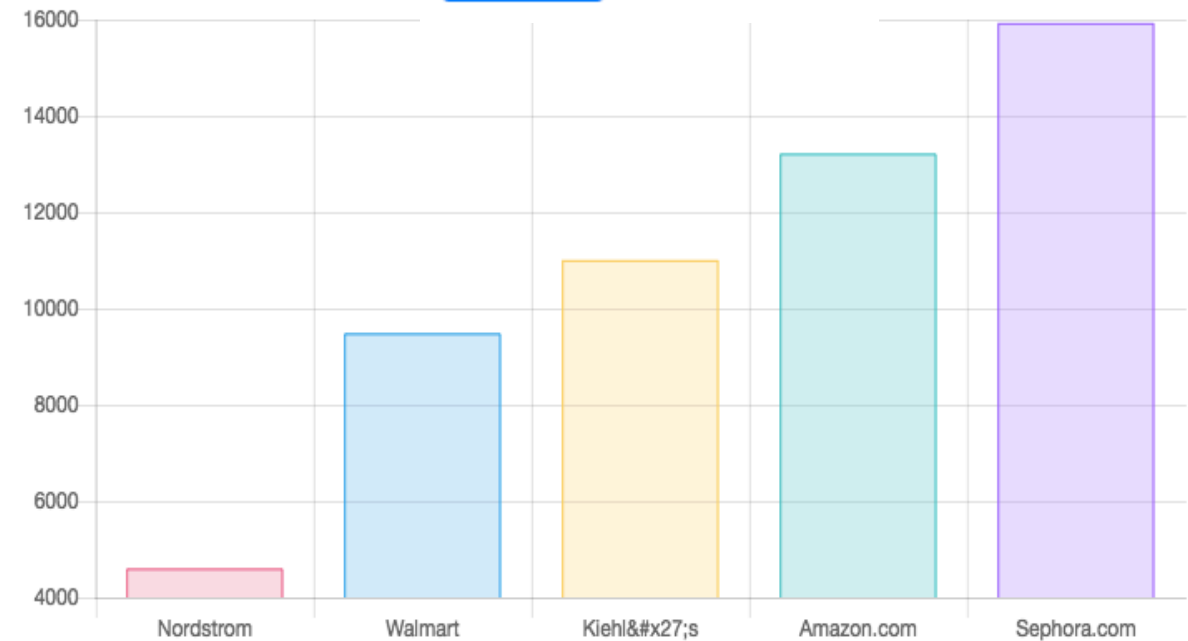
#### Top 5 Most Popular Platforms by Categories

Top 5 Popular Platforms:

category

Beauty

Submit



## Unit Test Cases

Word Spec Like  
format

```
"An master actor" must {  
  
  "start crawling and create the same number of child actors as pages" in {  
    val probe = TestProbe()  
    probe.send(master, CrawlerStart("https://www.dealmoon.com/Beauty", "Beauty", 20))  
    probe.expectMsg( obj = "Crawler Started!")  
    probe.expectMsg( obj = 20)  
  }  
}
```

Actor Test

*Based on Messages Transferring*

```
[info] ActorSpec:  
[info] An master actor  
[info] - must start crawling and create the same number of child actors as pages  
[info] - must Re-assign worker actors with Detail Page task and modify the crawled lists  
[info] An worker actor  
[info] - must receive parse main page job, return the urls list and finish signal  
[info] - must receive parse detail page job, return the discounts map and finish signal
```

Actor

Akka Test kit

Controller Test

*Based on HTML pages rendering*

```
[info] PlayControllerSpec:  
[info] Main Page  
[info] - should should have result of status code 200  
[info] Search most popular platform by category page  
[info] - should should have result of status code 200  
[info] Sort discounted product page  
[info] - should should have result of status code 200  
[info] Deal information statistics page
```

Controller

Scala Play Test

Test Result

```
[info] ScalaTest  
[info] Run completed in 45 seconds, 527 milliseconds.  
[info] Total number of tests run: 8  
[info] Suites: completed 2, aborted 0  
[info] Tests: succeeded 8, failed 0, canceled 0, ignored 0, pending 0  
[info] All tests passed.  
[info] Passed: Total 8, Failed 0, Errors 0, Passed 8  
[success] Total time: 59 s, completed Dec 13, 2018 1:19:58 PM
```



PART 3

**It is Demo Time !**



THANKS FOR  
WATCHING!

