LinkedIn Job Posting Data Analysis

Analysis By Vivek Chauhan

28-04-2025

Agenda

01. Data-Understanding & Data Cleaning

02. Uni-Variate Analysis

03. Bi-Variate Analysis

04. Multi-Variate Analysis

05. Insights/Recommendations

Introduction

• **Dataset Overview**: 123,849 rows × 31 columns of LinkedIn job postings data.

• **Objective**: Extract insights on job titles, salary trends, work types, and application methods.

• **Purpose**: Help job seekers and recruiters make informed decisions.

• **Methodology**: Data cleaning, transformation, and visualizations for trend identification.

	job_id	company_name	title	description
0	921716	Corcoran Sawyer Smith	Marketing Coordinator	Job descriptionA leading real estate firm in N...
1	1829192	NaN	Mental Health Therapist/Counselor	At Aspen Therapy and Wellness , we are committ...
2	10998357	The National Exemplar	Assitant Restaurant Manager	The National Exemplar is accepting application...
3	23221523	Abrams Fensterman, LLP	Senior Elder Law / Trusts and Estates Associat...	Senior Associate Attorney - Elder Law / Trusts...
4	35982263	NaN	Service Technician	Looking for HVAC service tech with experience ...
...	...	...	...	...
123844	3906267117	Lozano Smith	Title IX/Investigations Attorney	Our Walnut Creek office is currently seeking a...

```
[53]:  # Let's count the words in title column so we get the common word in a job position

       from collections import Counter # load the counter library
       from nltk.corpus import stopwords  # load the stopwords library
       import nltk

       nltk.download('stopwords')  # download stopwords (only needed once)

       stop_words = set(stopwords.words('english'))  # load English stopwords

       all_title_words = [] # create a blank list to store split words in it

       for title in new_data['title']:     # loop for split the words and store it in a blank list
           words = title.split()
           filtered_words = [word for word in words if word.lower() not in stop_words]  # remove stopwords
           all_title_words.extend(filtered_words)  # add filtered words to the list

       title_word_freq = Counter(all_title_words)  # count the split words

       print(title_word_freq.most_common(10))  # print the 10 most common words
```

```
[nltk_data] Downloading package stopwords to C:\Users\vivek
[nltk_data]      chauhan\AppData\Roaming\nltk_data...
[nltk_data]   Package stopwords is already up-to-date!
[('-', 31473), ('Manager', 15087), ('Engineer', 9063), ('Sales', 8031), ('Senior', 7898), ('Specialist', 6140), ('Associate', 5961), ('Assistant', 570
9), ('Technician', 4911), ('Analyst', 4350)]
```

Top 10 most common words in the job title column is 'Manager','Engineer','Sales','Senior','Specialist','Associate','Assistant','Technician','Analyst'.

```
[55]:   # let's count the words in description column so we get the common word in a job description
        # we get better idea to create ATS friendly resume for applying job position

        from nltk.corpus import stopwords  # import stopwords
        import nltk

        stop_words = set(stopwords.words('english'))  # load English stopwords

        all_description_words = []  # create a blank list to store split words

        for description in new_data['description']:  # loop for splitting the words
            words = description.split()
            filtered_words = [word for word in words if word.lower() not in stop_words]  # remove stopwords
            all_description_words.extend(filtered_words)  # add filtered words to the list

        description_word_freq = Counter(all_description_words)  # count the split words

        print(description_word_freq.most_common(25))  # print the 25
```

[('work', 257795), ('experience', 210779), ('team', 167754), ('including', 153531), ('business', 124584), ('customer', 112487), ('&', 110836), ('-', 107962), ('support', 104537), ('years', 103156), ('may', 95263), ('management', 89380), ('position', 88447), ('skills', 88330), ('ability', 85159), ('new', 84903), ('working', 83743), ('within', 82518), ('required', 82009), ('care', 81704), ('company', 81671), ('related', 81093), ('sales', 79869), ('and/or', 77765), ('data', 77626)]

In the company job discription the common words 'work', 'experience', 'team', 'including', 'business', 'customer', 'support', 'years', 'management', 'position', 'skills', 'ability', 'sales' etc.

5

```python
[56]:  # Let's count the words in location column so we get the common location in a job position

       from collections import Counter # load the counter library
       from nltk.corpus import stopwords  # import stopwords
       import nltk

       nltk.download('stopwords')  # download stopwords (only needed once)

       stop_words = set(stopwords.words('english'))  # load English stopwords

       all_location_words = []  # create a blank list to store split words

       for location in new_data['location']:  # loop for splitting the words
           words = location.split()
           filtered_words = [word for word in words if word.lower() not in stop_words]  # remove stopwords
           all_location_words.extend(filtered_words)  # add filtered words to the list

       location_word_freq = Counter(all_location_words)  # count the split words

       print(location_word_freq.most_common(10))  # print the 10 most common words
```

```
[nltk_data] Downloading package stopwords to C:\Users\vivek
[nltk_data]     chauhan\AppData\Roaming\nltk_data...
[nltk_data]   Package stopwords is already up-to-date!
[('United', 12695), ('States', 12695), ('CA', 11484), ('TX', 10271), ('NY', 6044), ('FL', 5907), ('New', 5561), ('NC', 4928), ('Area', 4810), ('IL', 4486)]
```

As you know that this dataset is contains only usa details so common sense locations is 'united states' but the few of top 3 job posting locations is 'CA','TX','NY' & 'FL','NC','IL'.

Pay Period Demand

Job Posting pay_period Category
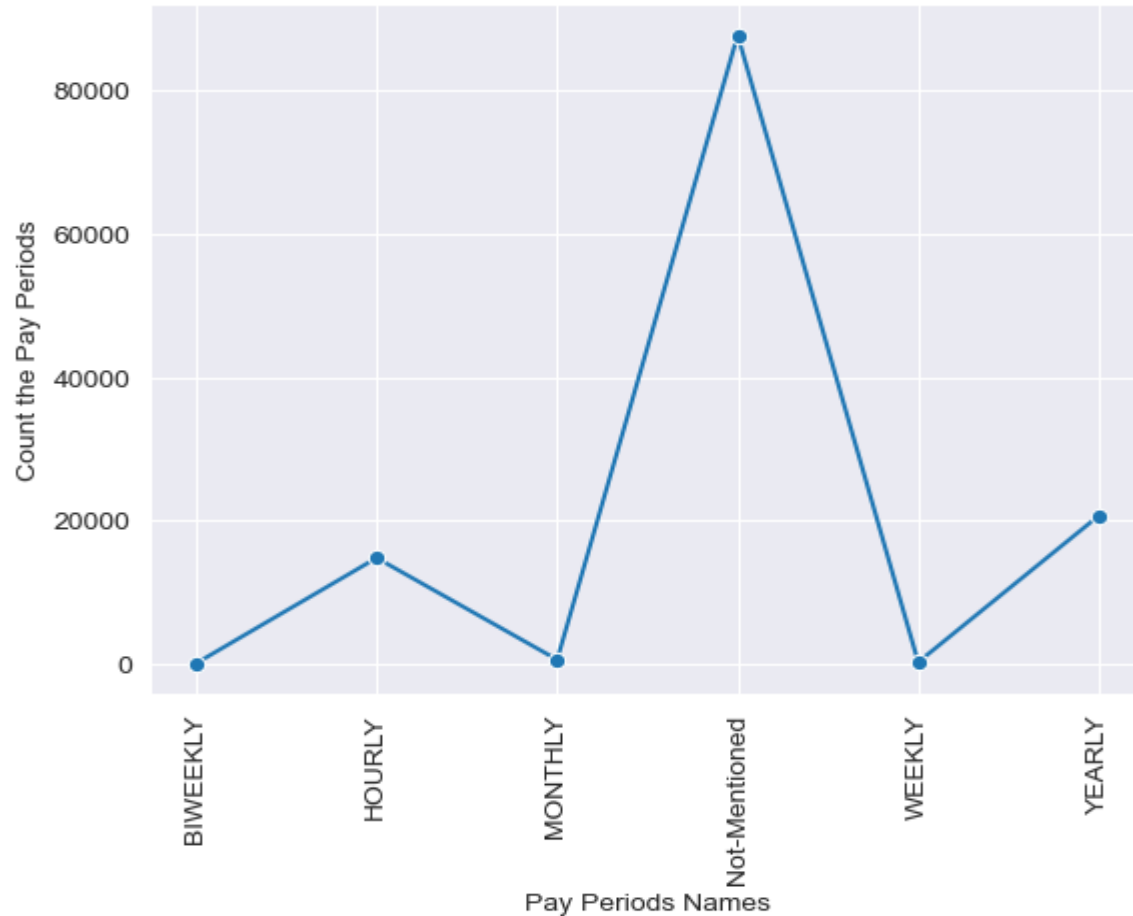
In the above chart as per the accurate data there is Yearly Pay types jobs is highest on linkedin and rest of the data we have no information so "Not-Mentioned Category" is high if data is proper then game is changed.
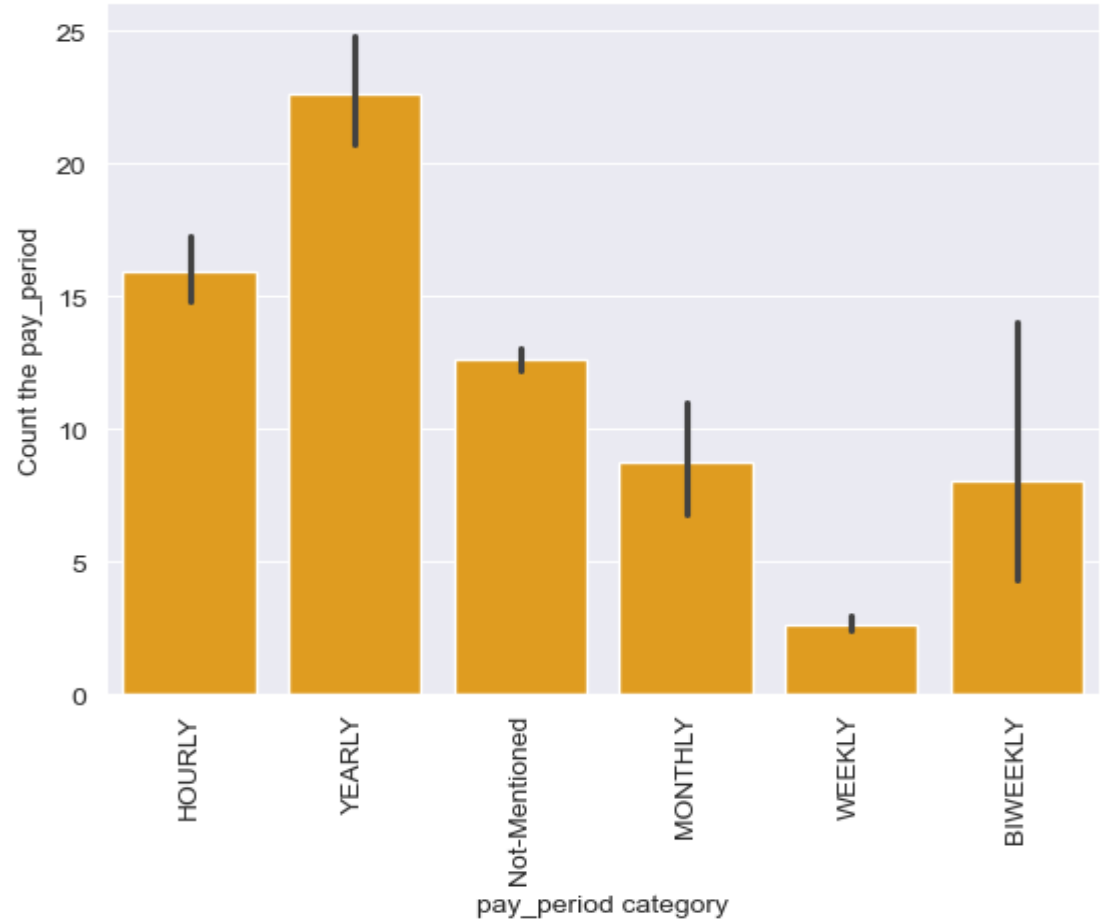
Most of the company offered the pay period category type is "Yearly" basis on the linkedin platform.

Job Posting views_cat_func
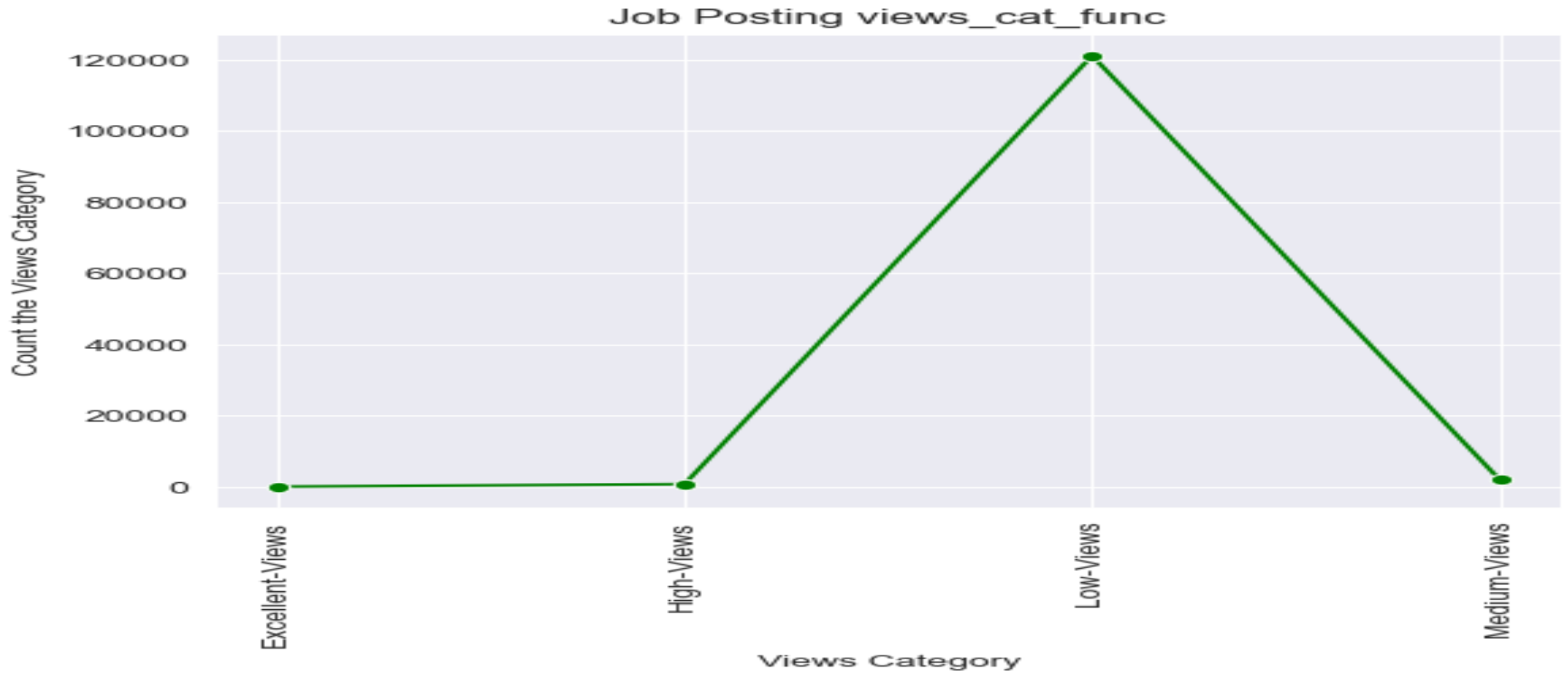
You can see clearly most of the jobs are posted on linkedin platform fall into low views category.

Job Posting applies_cat_func Category
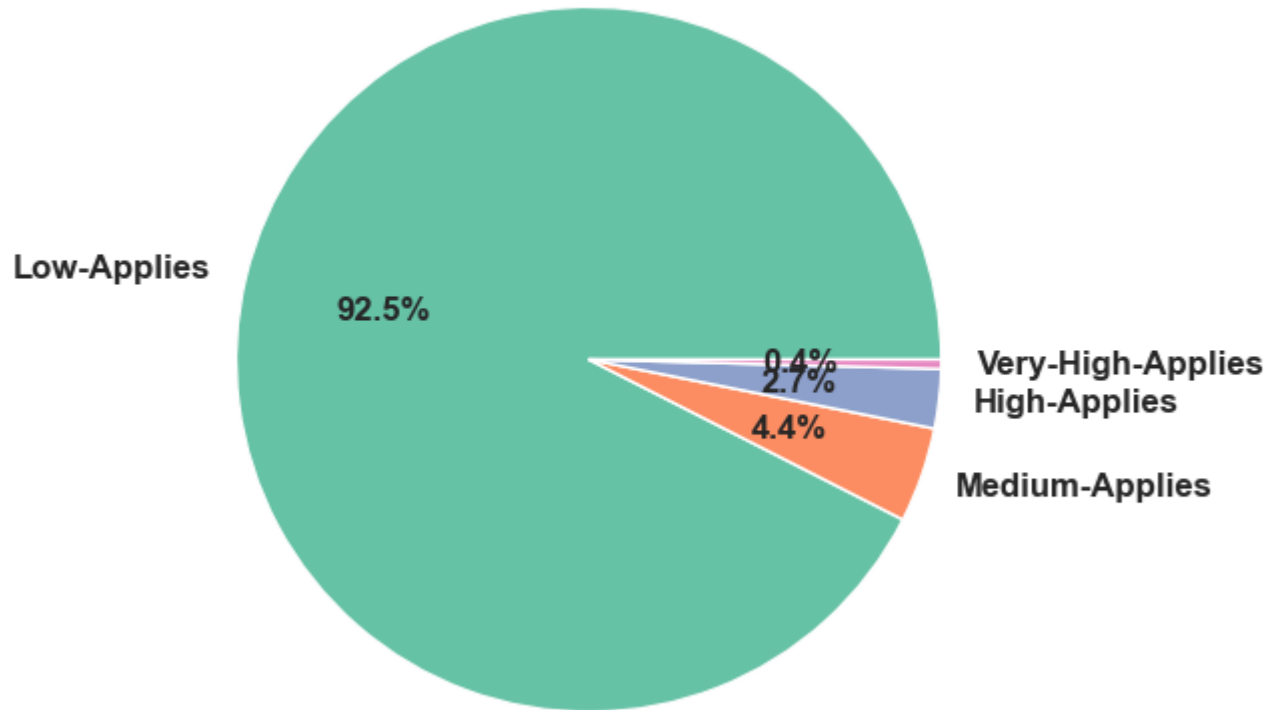
Most of the job posting are fall into Low-Applies category.
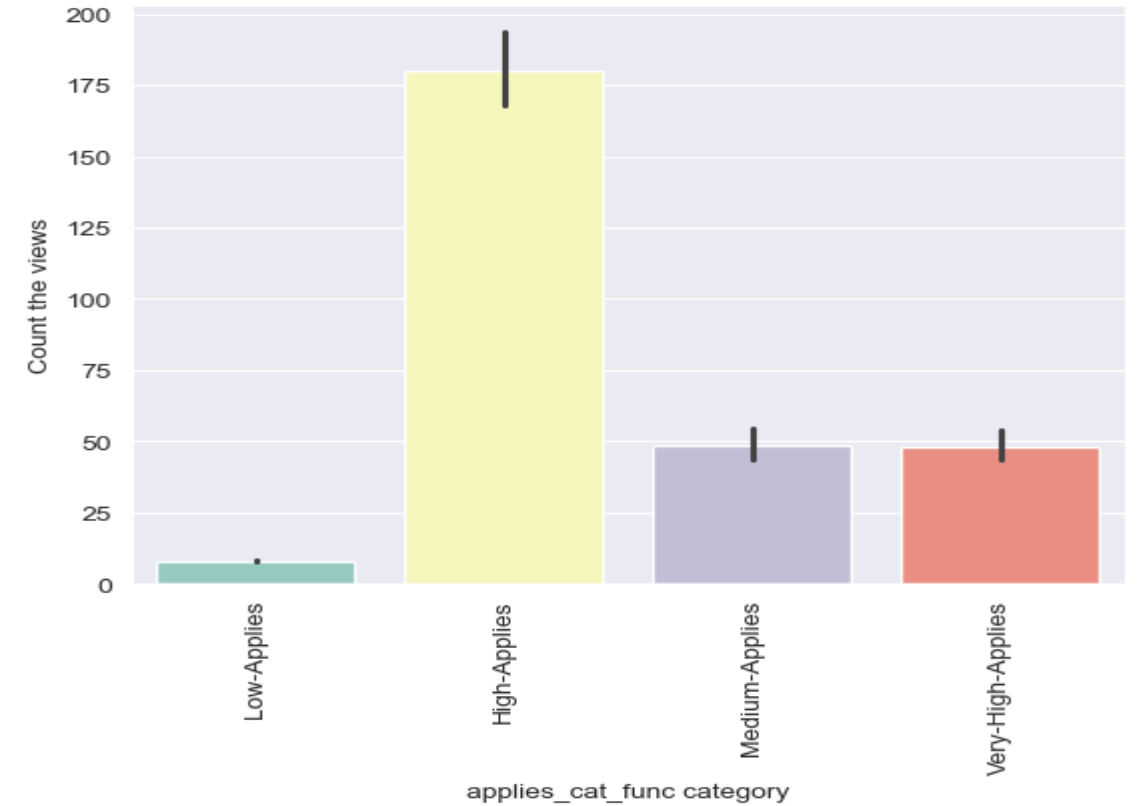
In the job posting applicants is high getting the highest views on it.

Job Posting min_salary_func Category
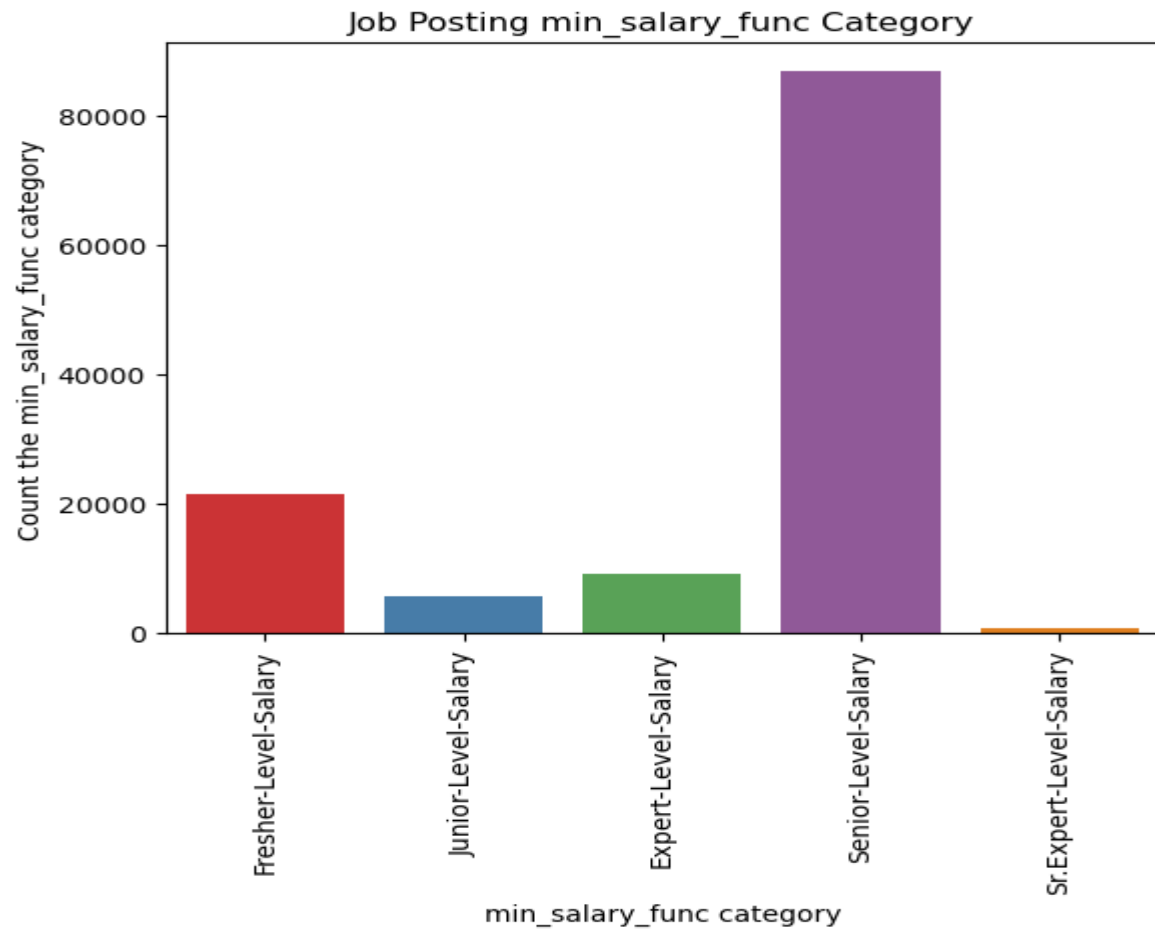
As you can see the 'Senior-Level-Salary' is high in the min-salary on linkedin plateform.
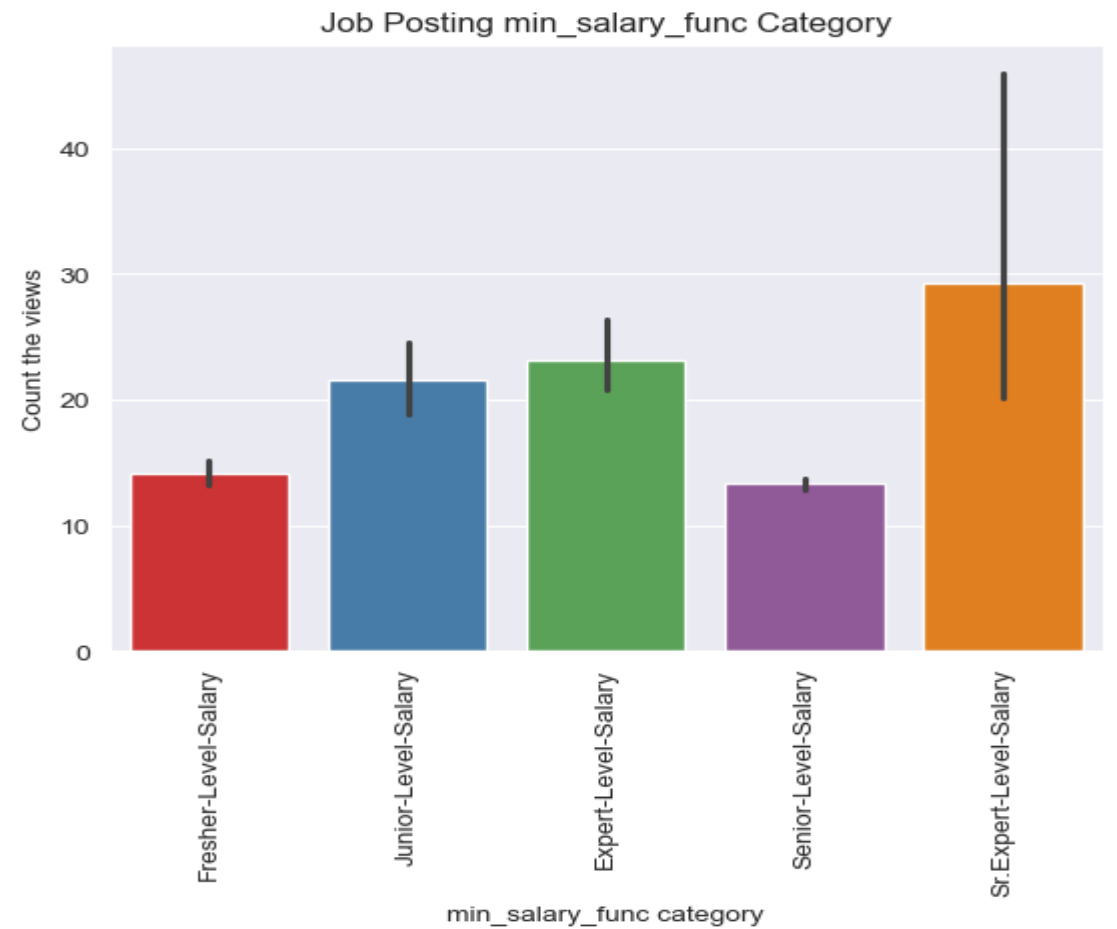
Those company offer the minimum salary in the category of 'Expert Level Salary' & 'Sr.Expert Level Salry' getting the highest views on linkedin platform.

Job Posting max_salary_cat_func Category

Job Posting max_salary_func Category
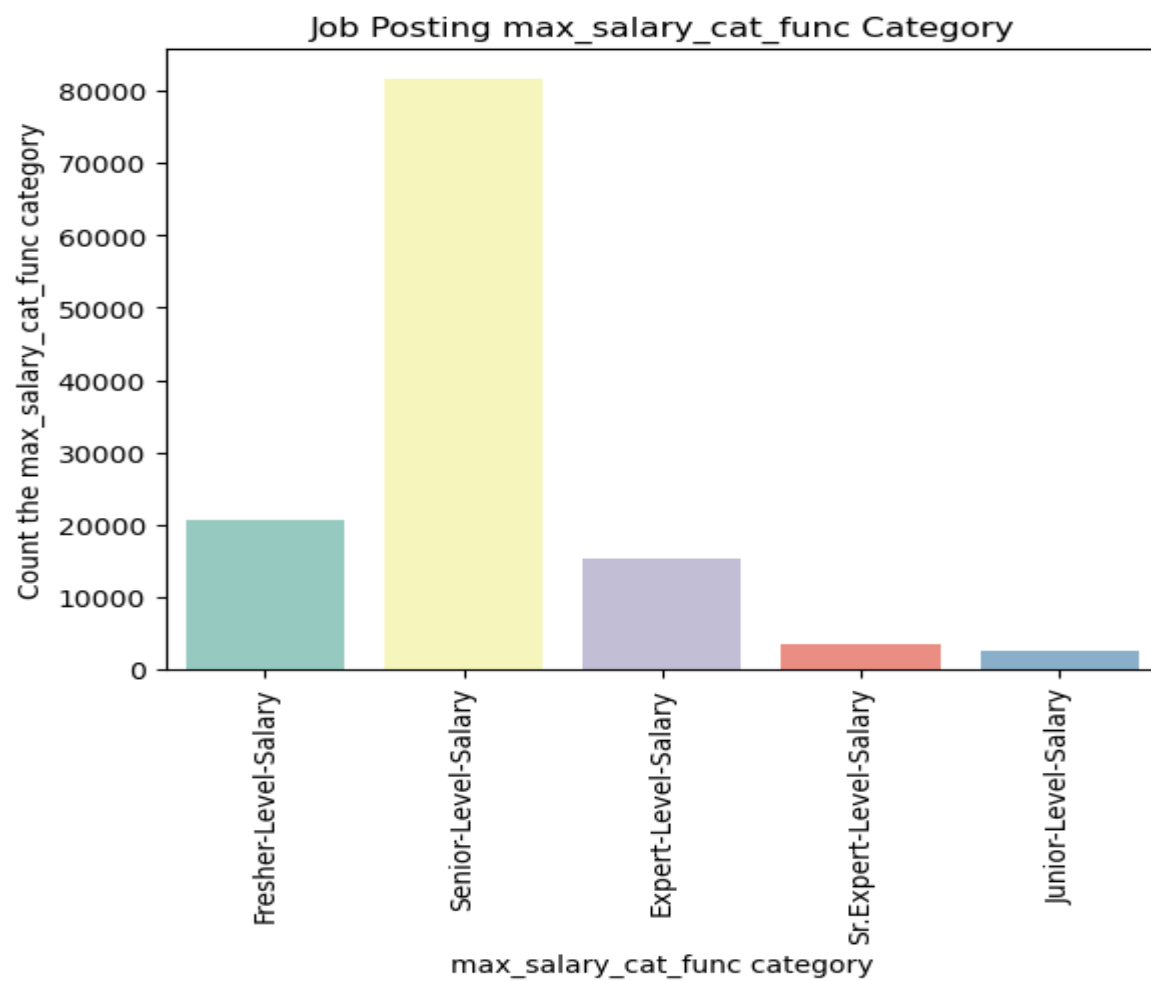
Again you can see the winner is 'Senior-Level-Salary' is high in the max_salary criteria on the linkedin platform.
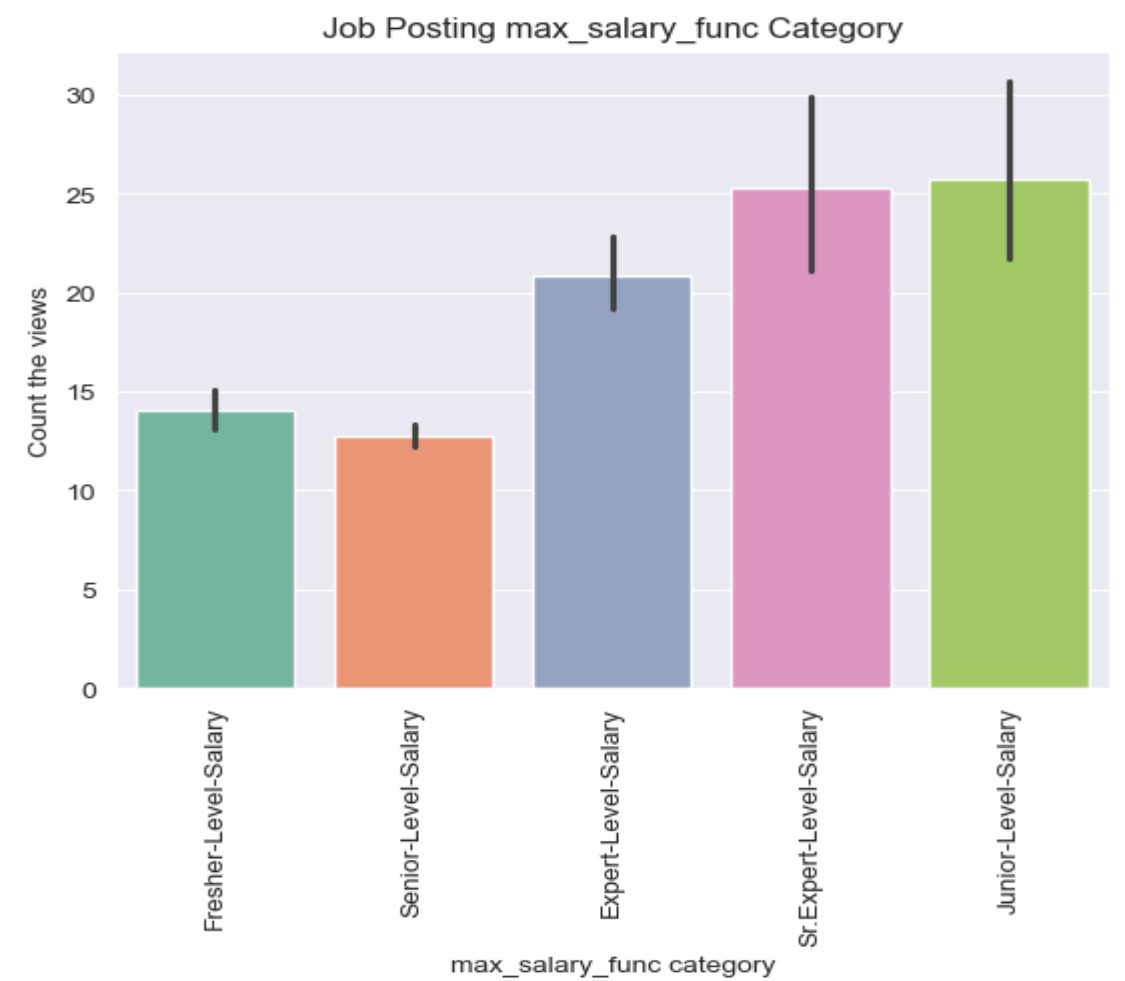
Those company offer the maximum salary in the category of 'Sr.Expert Level' & 'Junior Level Salary' getting the highest views.
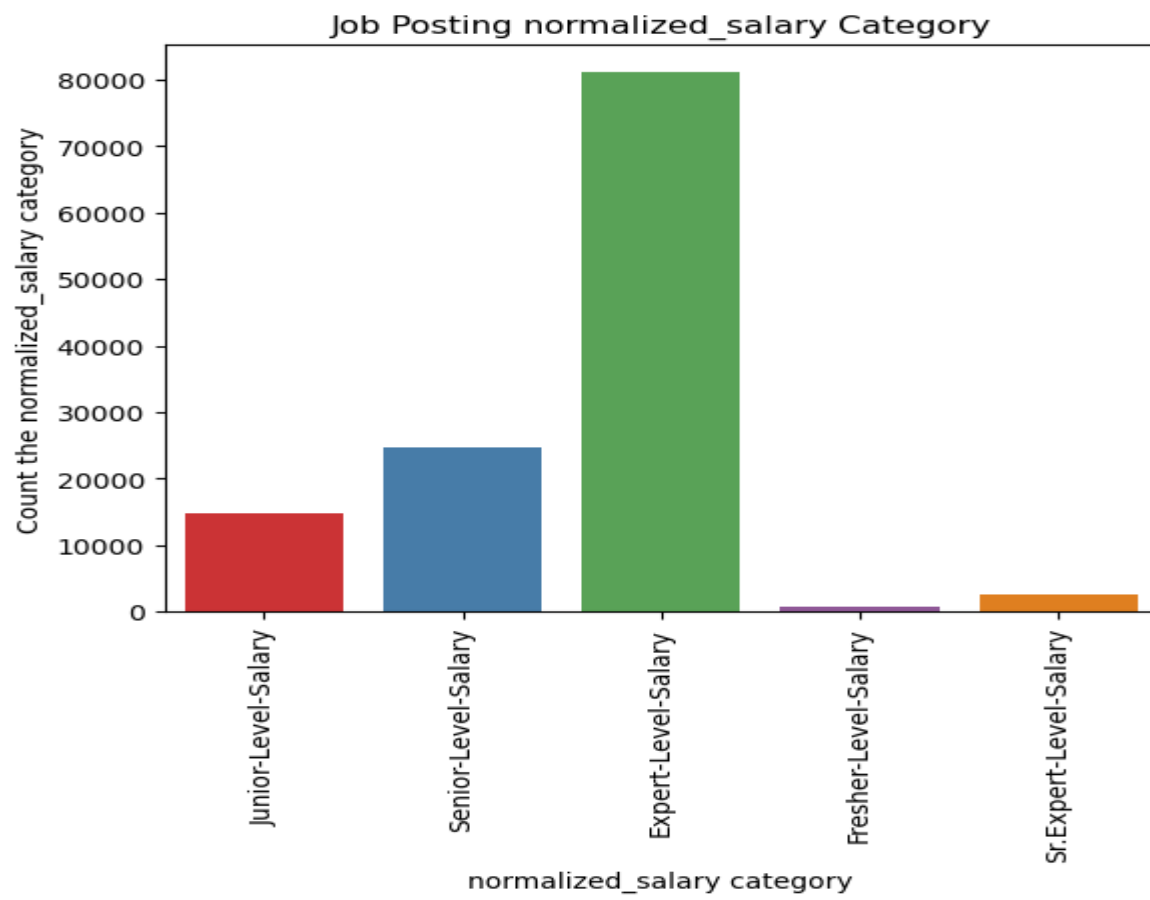
You can see clearly we have not proper data to analyze it but as per this data "Expert-level-Salary" in normalized salary category is most offered by company.

Those company offers the normalized salary in the category of 'Sr.Expert Level' getting the highest views.

Job Posting formatted_work_type Category
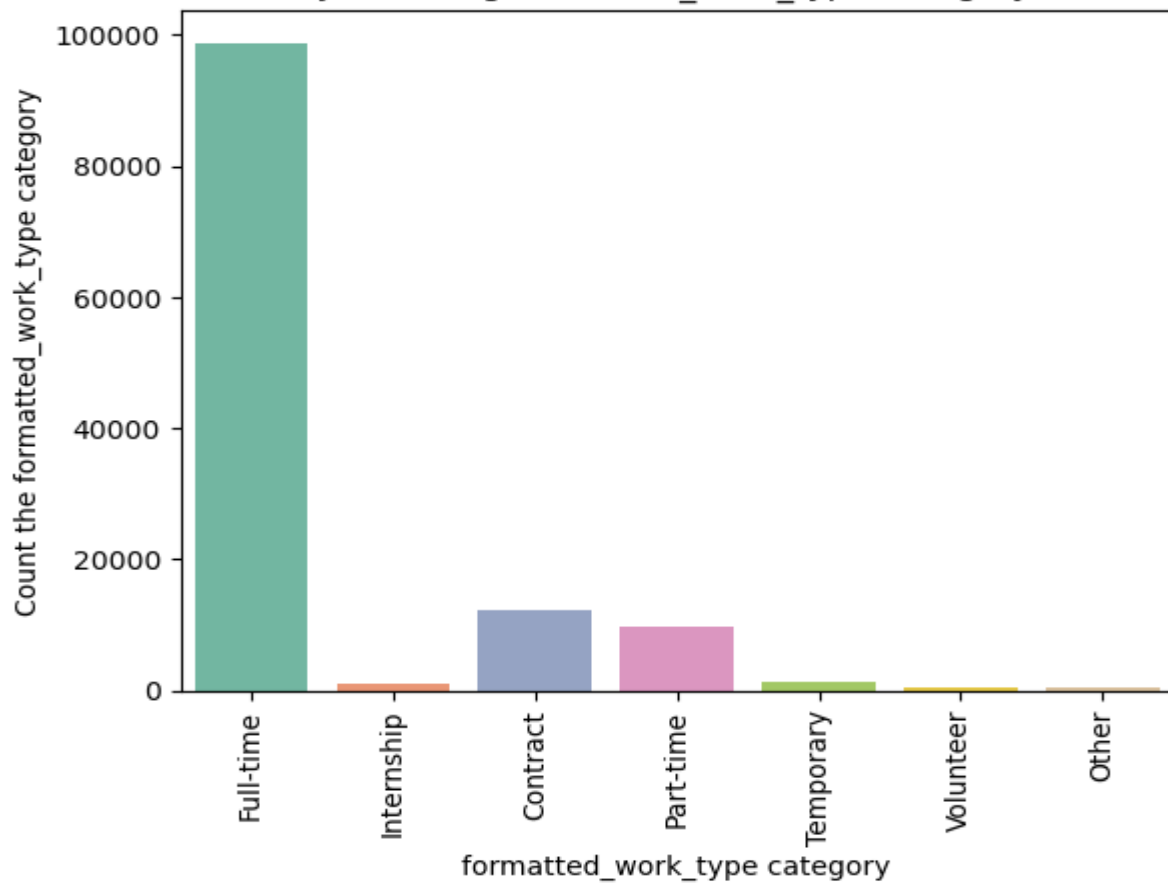
Most of the jobs work type is Full-time.
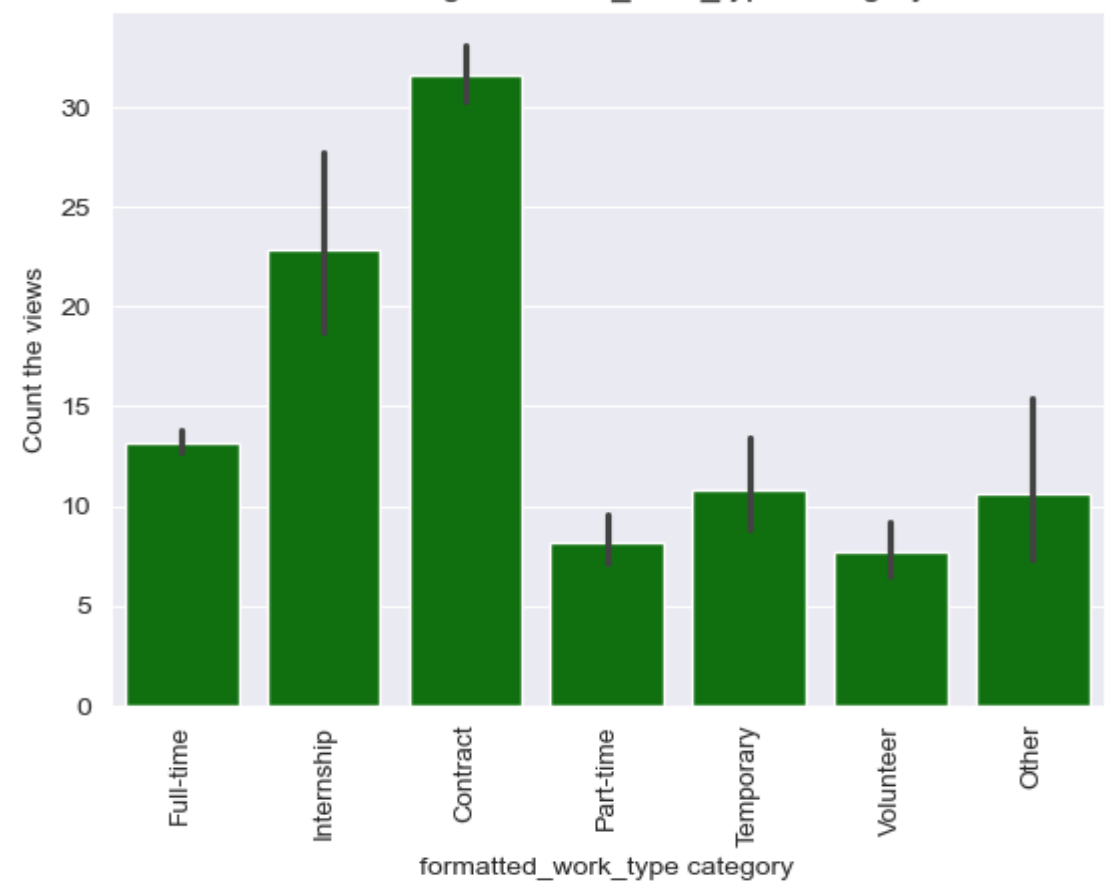
Most of the job posting offered the formatted work type category on 'Contract' bases on linkedin platform.

Job Posting application_type Category

Job Posting application_type Category
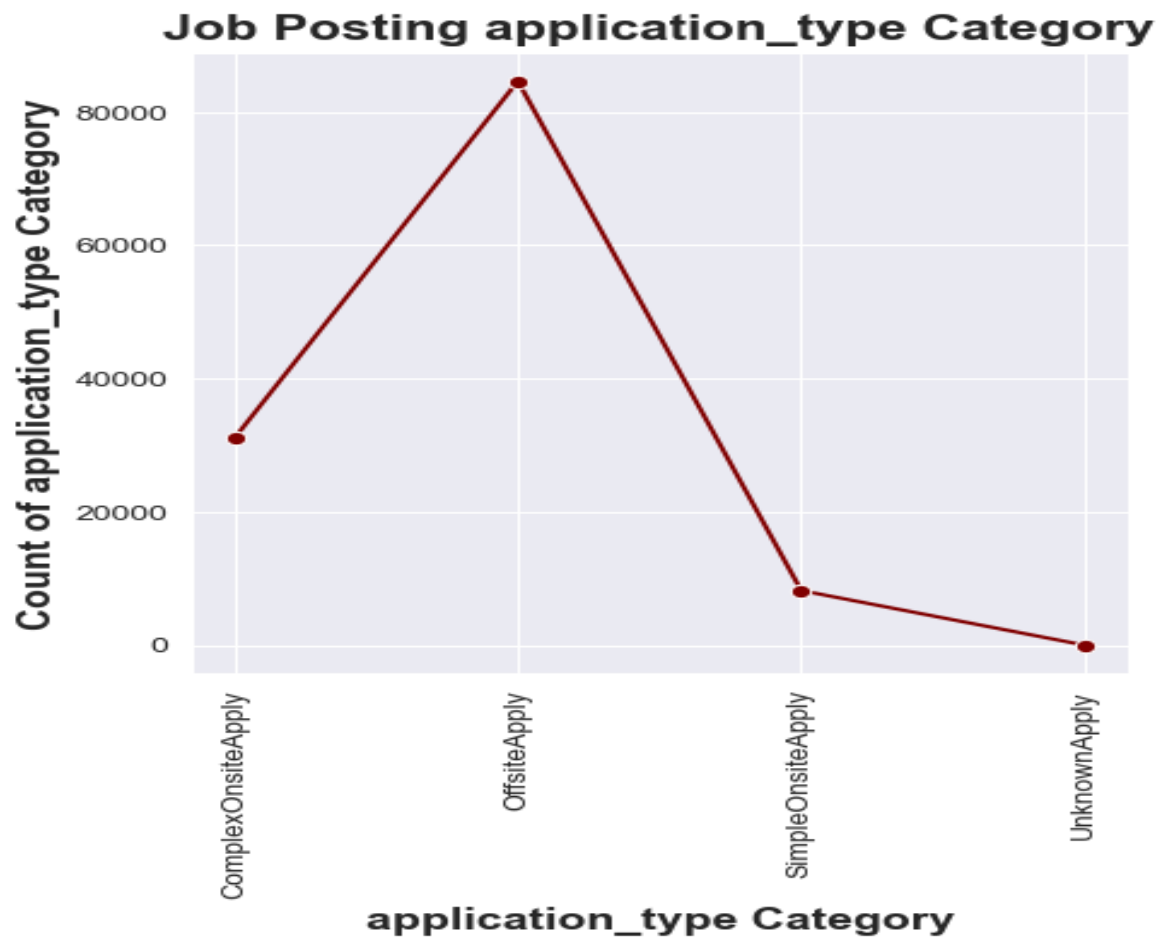
In the most of the jobs which is posted on linkedin platform application type is "OffsiteApply" is high.
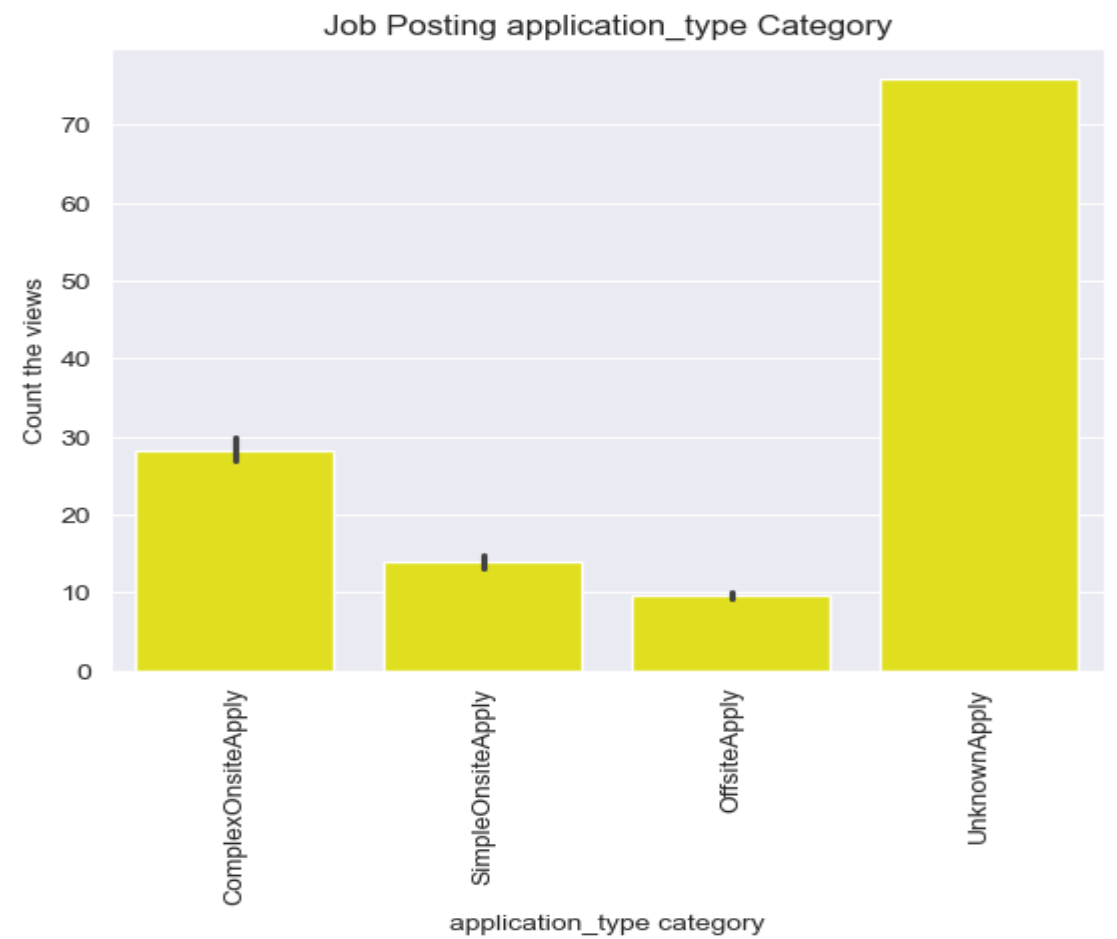
Those who posted the jobs in 'Unknown Apply' category getting the highest views on linkedin platform.

Job Posting formatted_experience_level Category

Job Posting formatted_experience_level Category
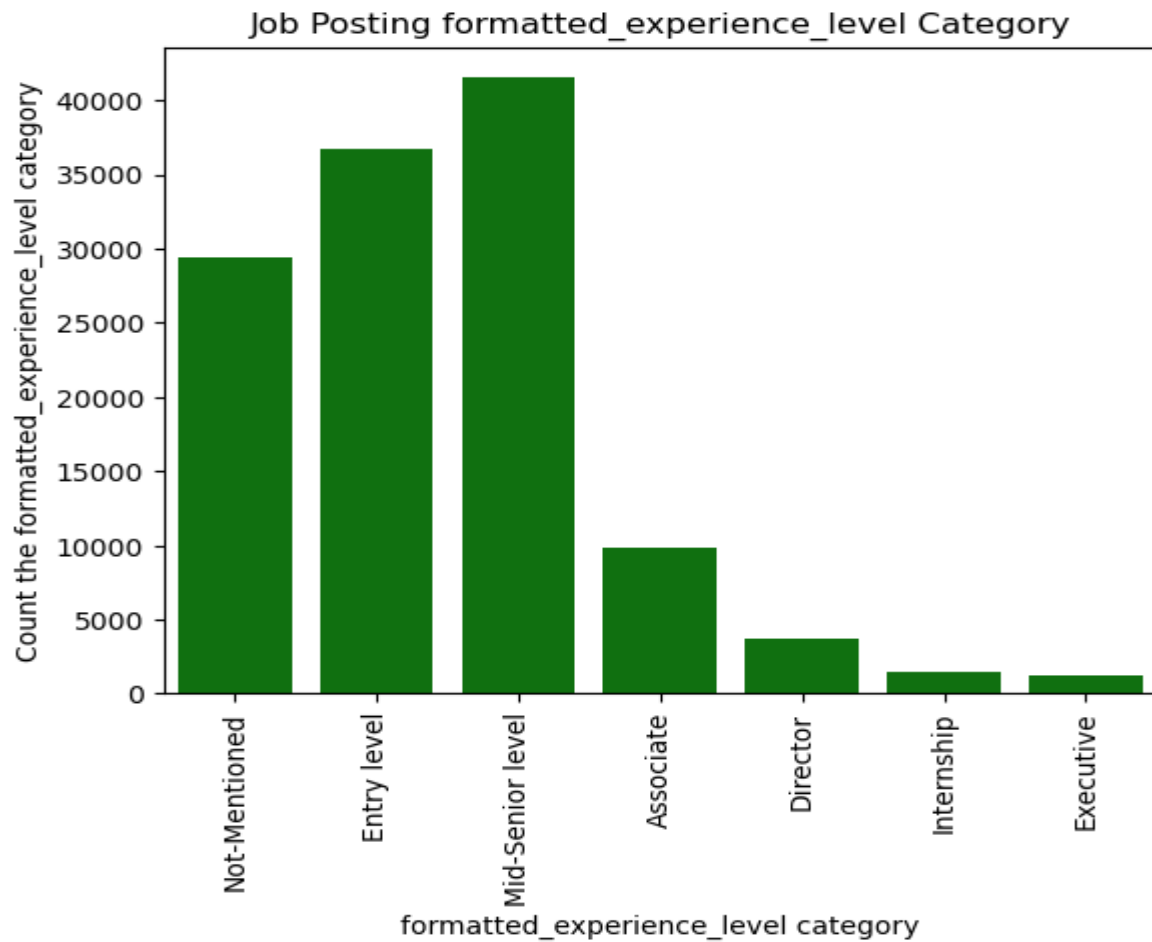
Most of the jobs posting requirement is "Mid-Senior level" & second position is "Entry-Level".
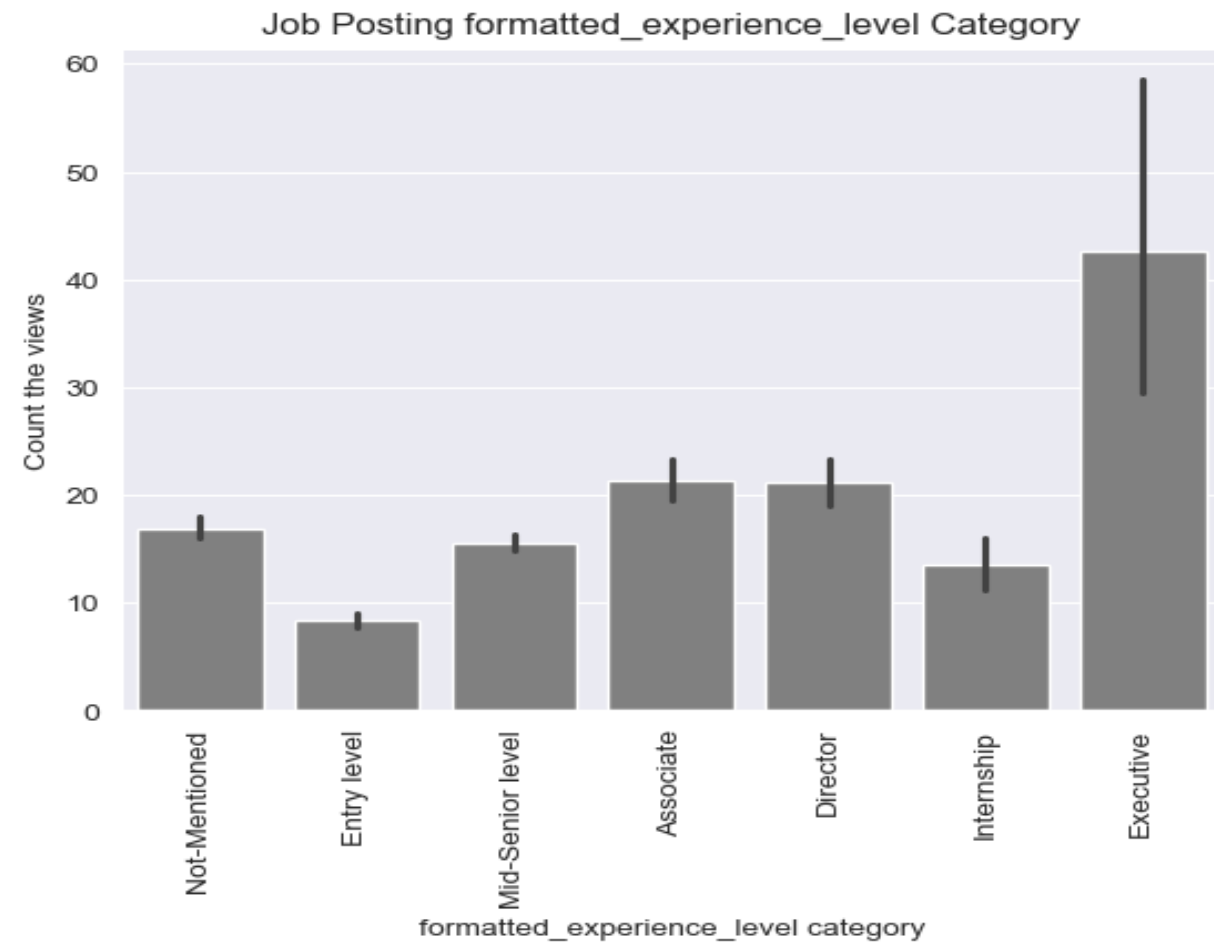
Here the highest variance in the views in 'Executive' formatted experience level but if we want accurate formatted experience level category then the winner is 'Associate' & 'Director' getting the highest views on linkedin platform.

15

Job Posting work_type Category
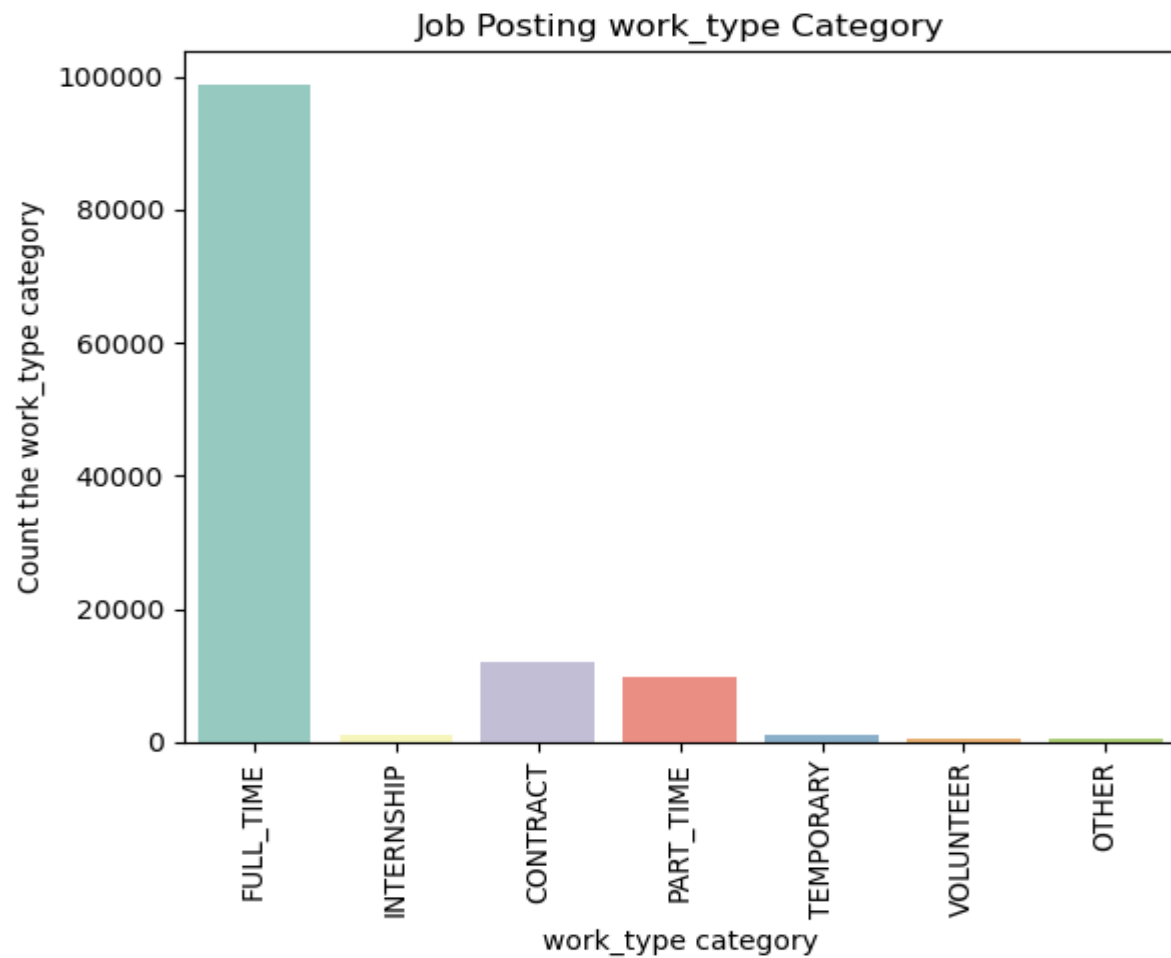
In above chart you can see clearly the work type category is "Full-Time".
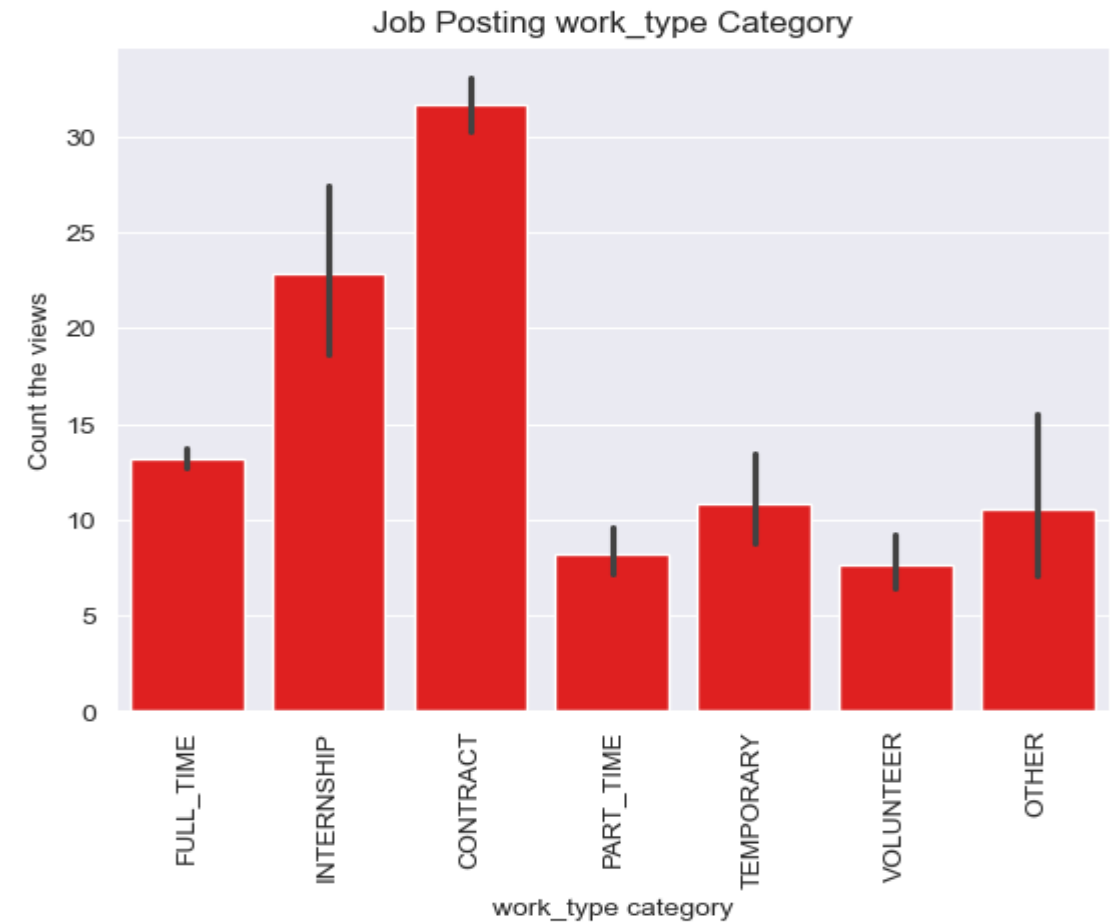
You can see clearly 'Contract' based work type cotegory getting the highest public views on linkedin platform.

Job Posting compensation_type Category
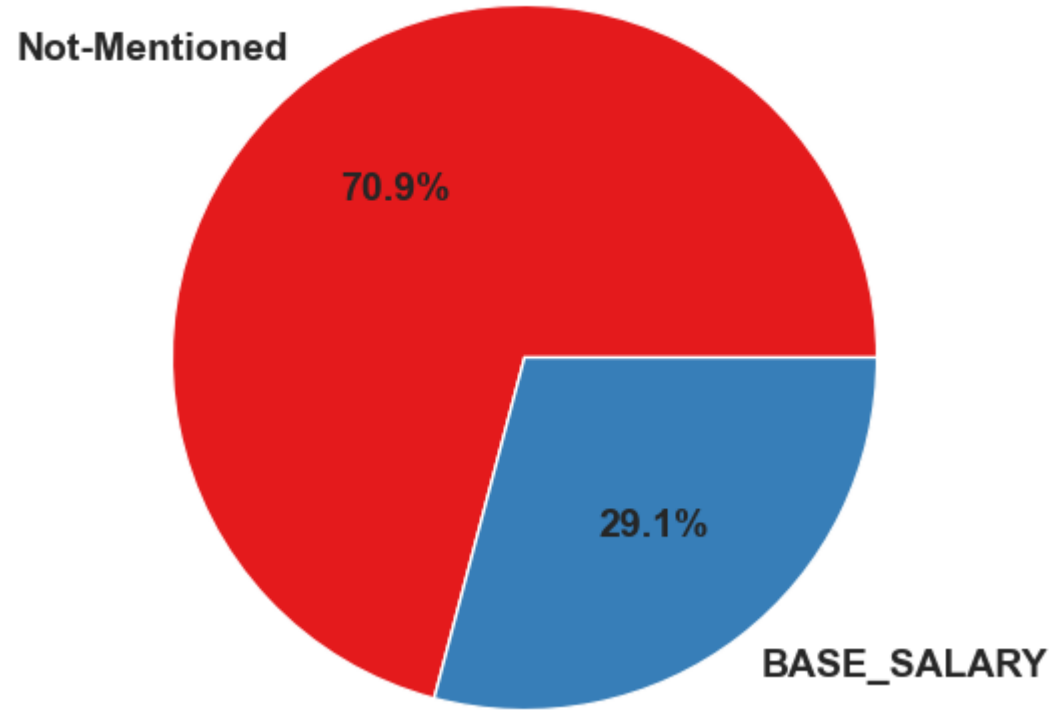
we have not proper or you can say accurate data that's we fill Not-Mentioned but if you consider raw data and analyze it Most of the "Base Salary" Compensetion Type is high.

As you know that we have not suffienct or proper data to analyse compensation type column but in the improper data the winner is 'Base Salary' compensation type category getting the highest views on linkedin platform.

Job Posting expmonth Category

Job Posting expmonth Category
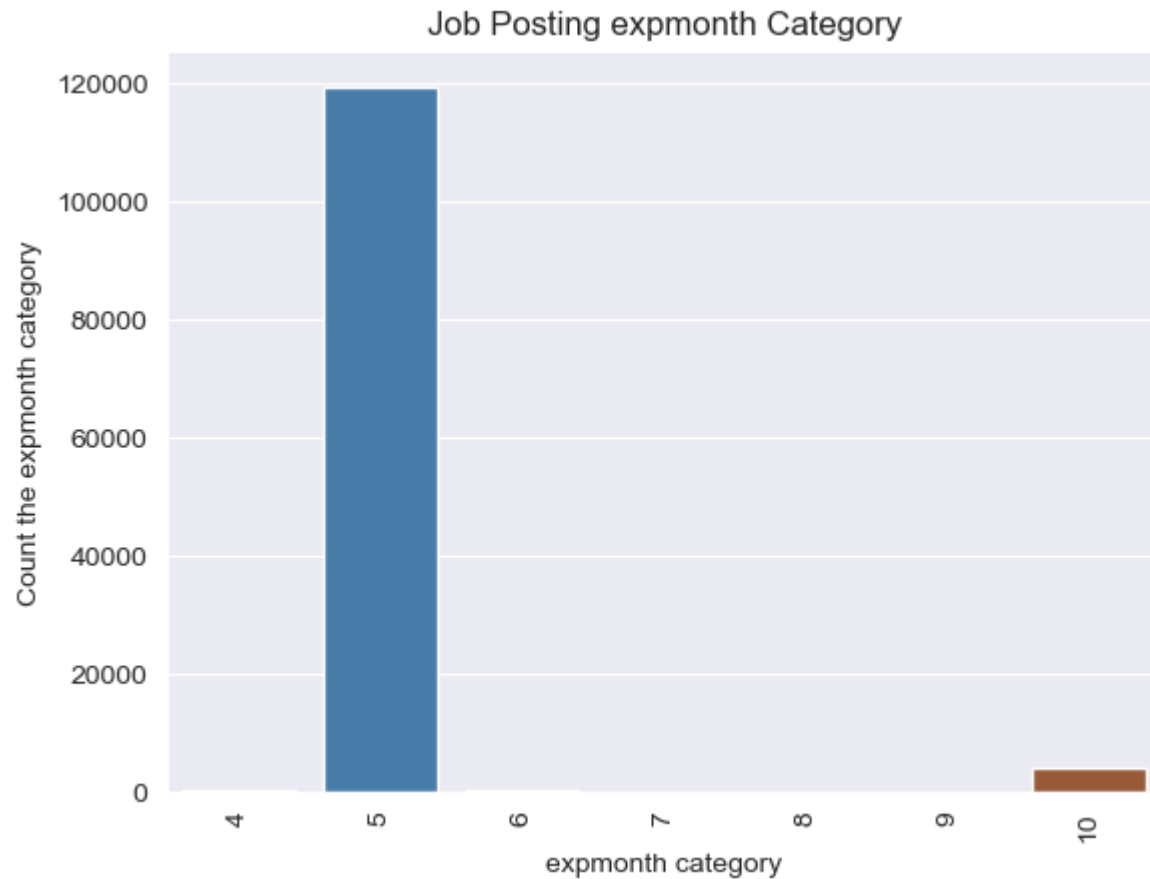
Most of the job-posting expiry month is May may be Applicants enjoy summer vacation joke a part.
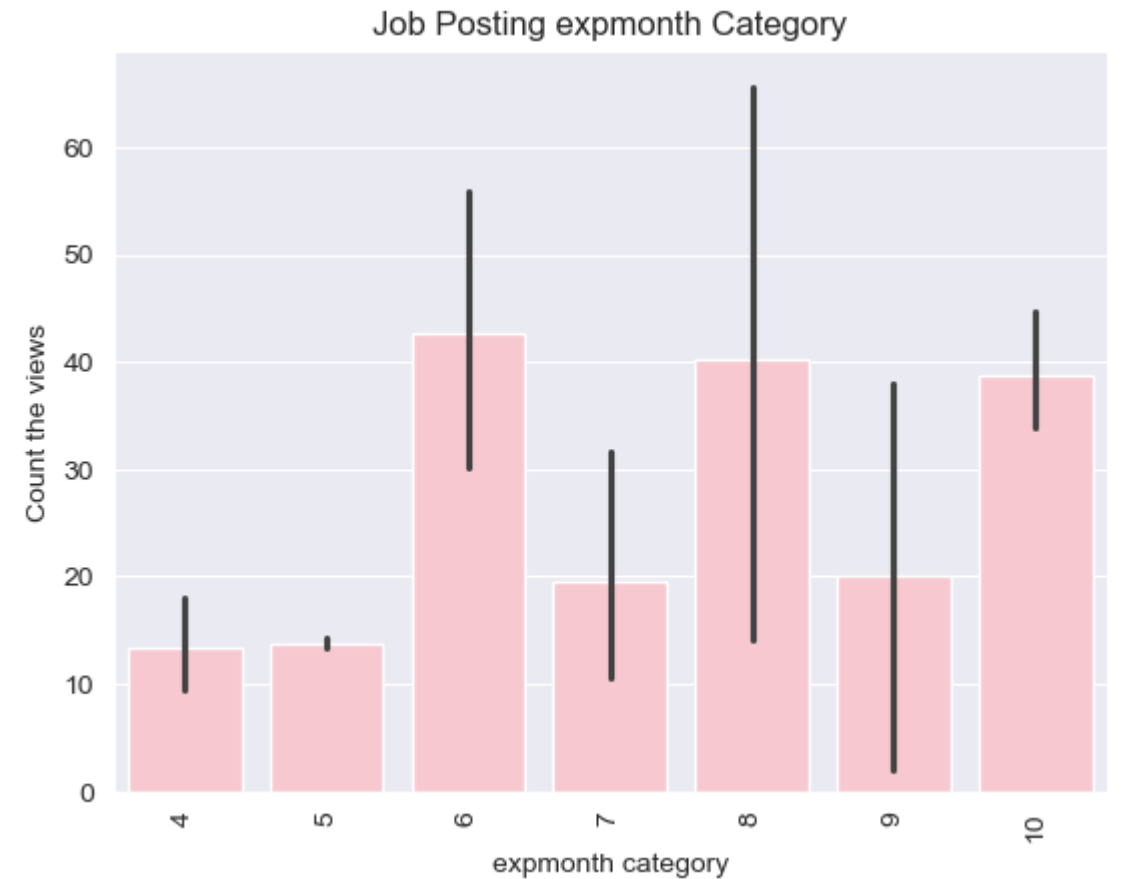
Job posting expiry date in the month of 'October' getting the highest views may be at last viewers is high.
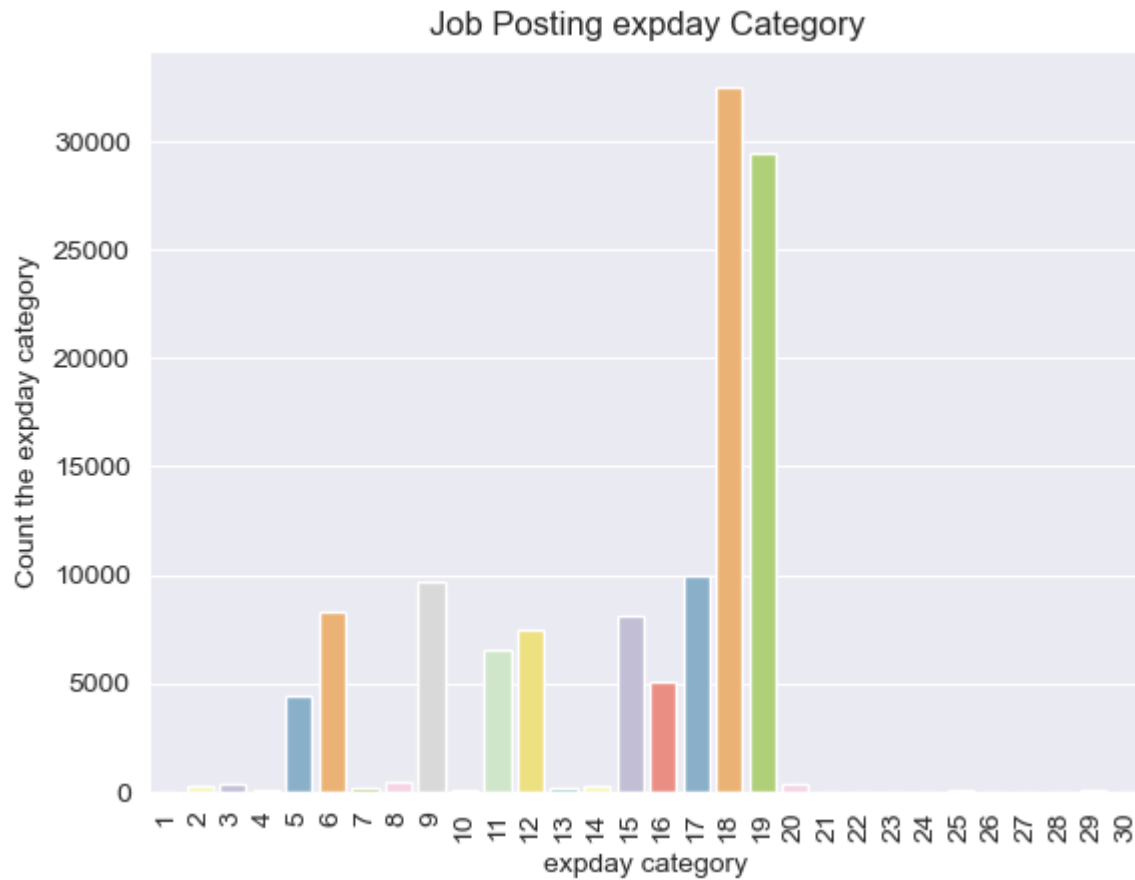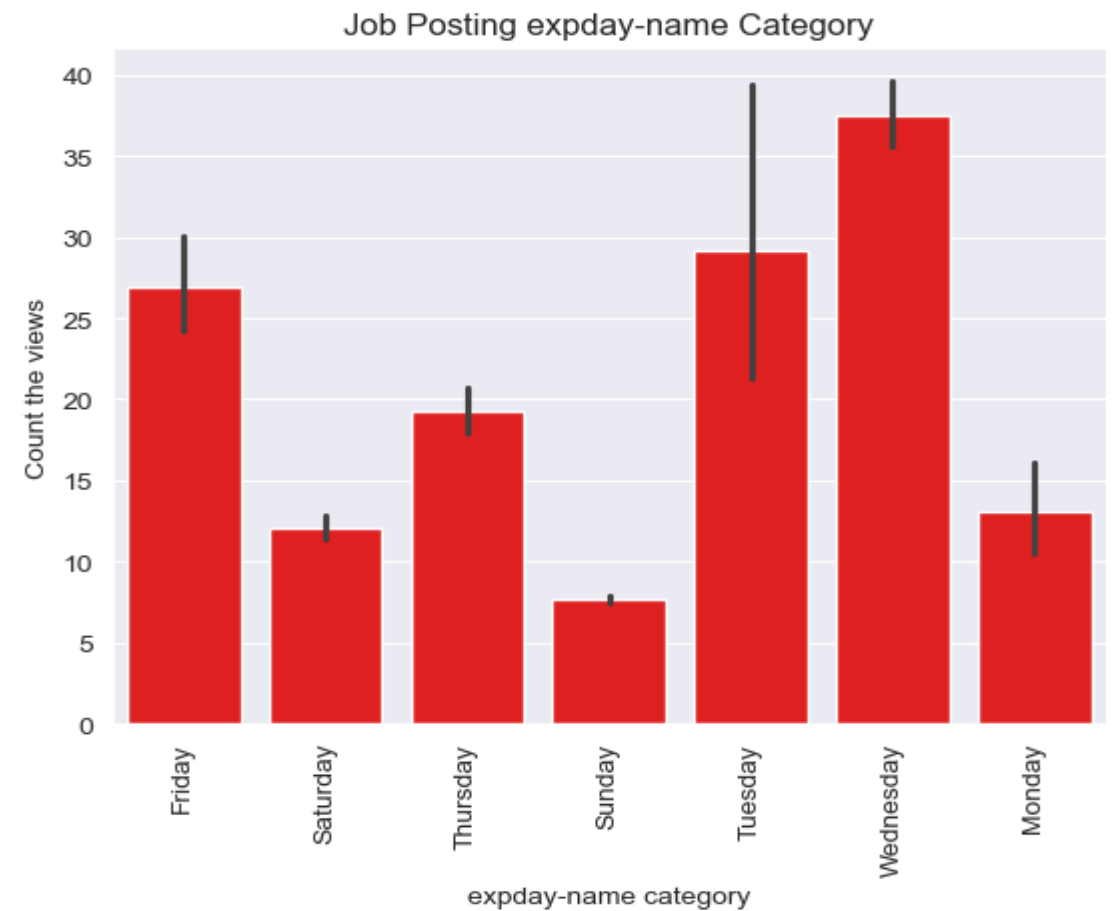
**Most of the job posting expired in the date of 18th May.**
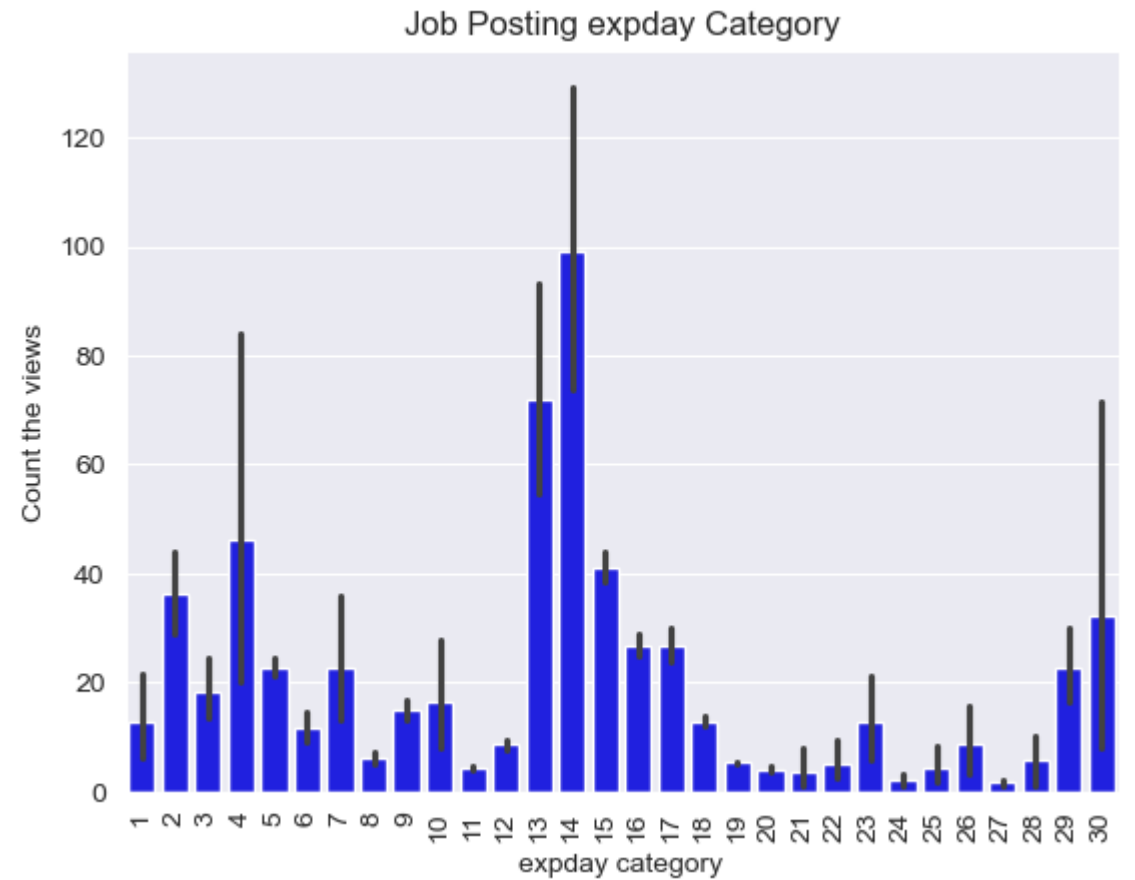
**In the expiry date of 13th getting the highest views on linkedin platform.**
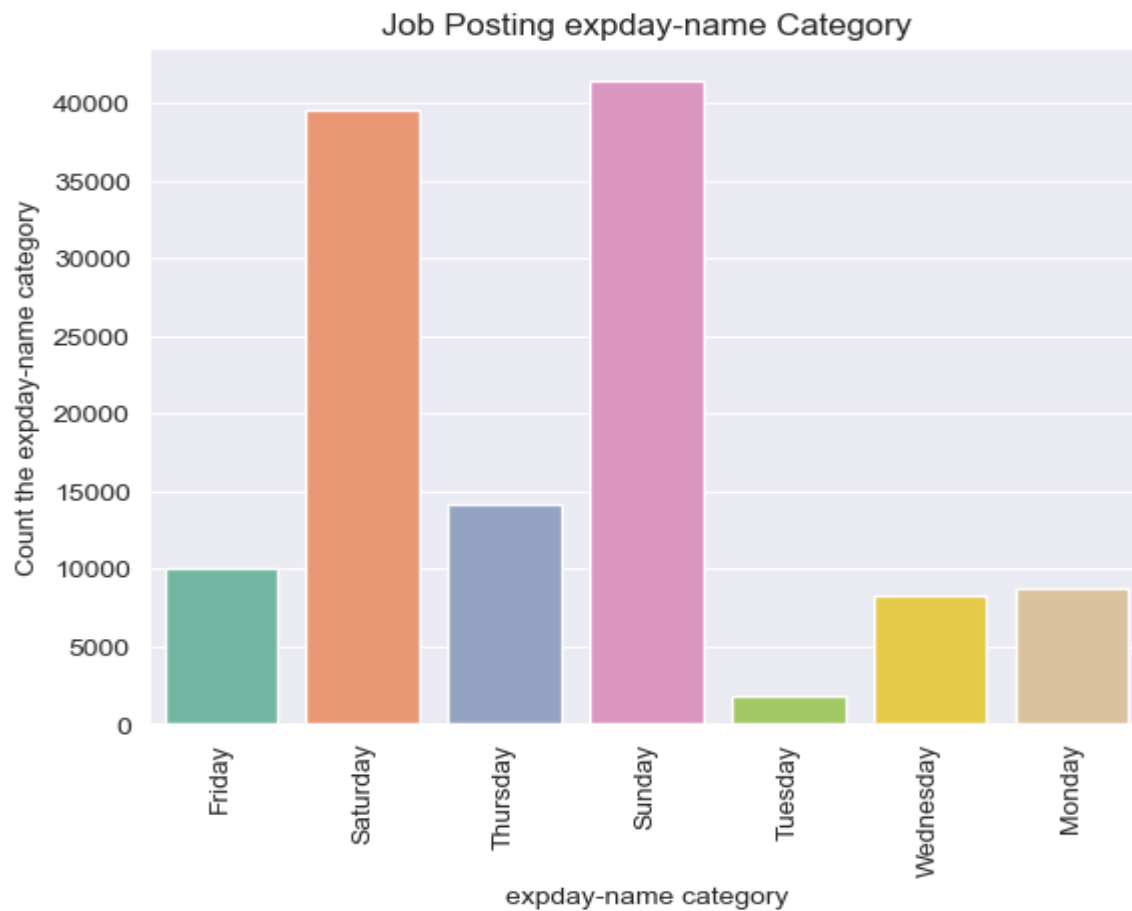
Most of the job posting are expired in the day of Sunday May be Applicants Enjoy Weekend.
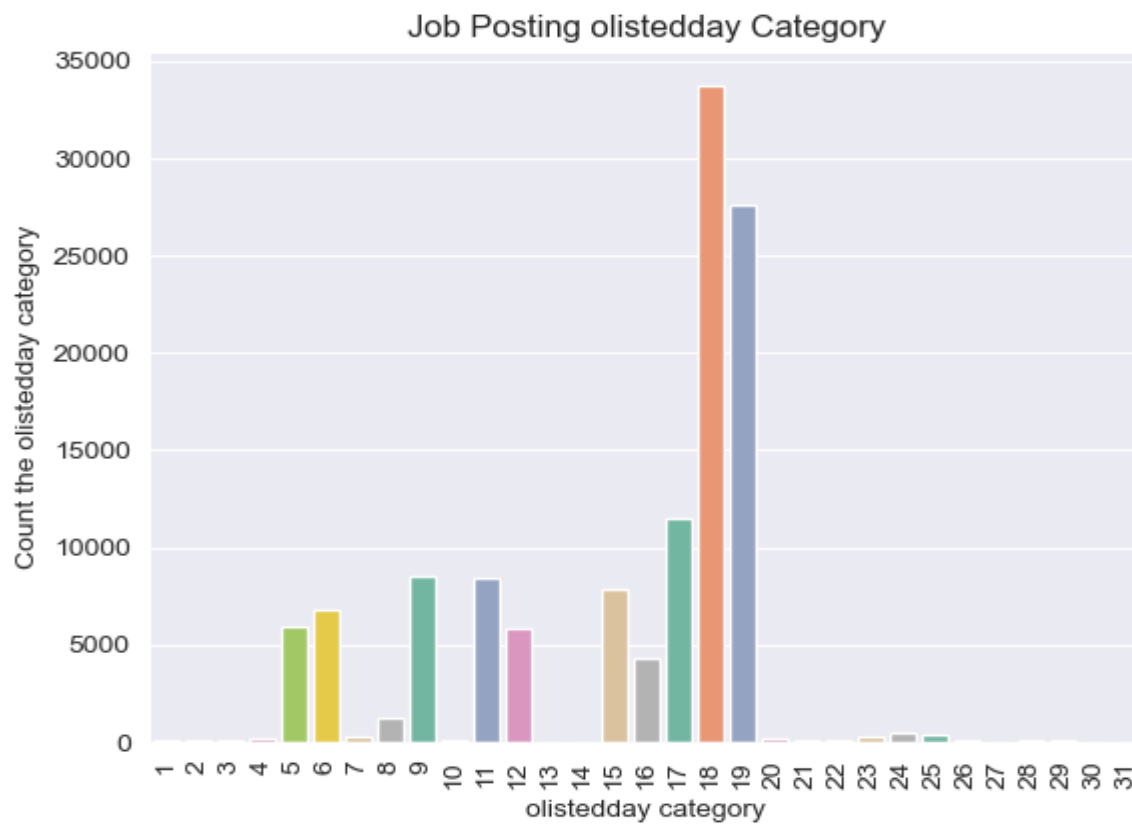
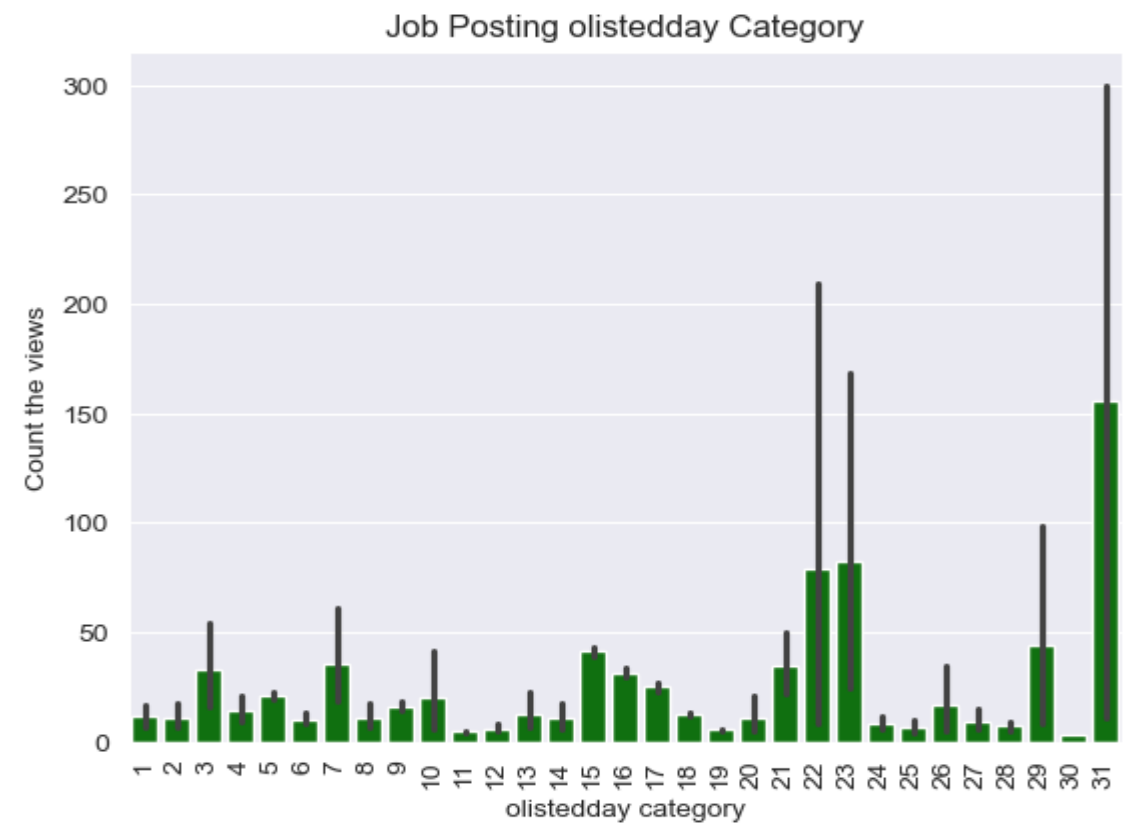**Those job posting expiry date in on the day of 'Wednesday' getting the highest views on linkedin platform.**

Job Posting olistedmonth Category

Job Posting olistedmonth Category
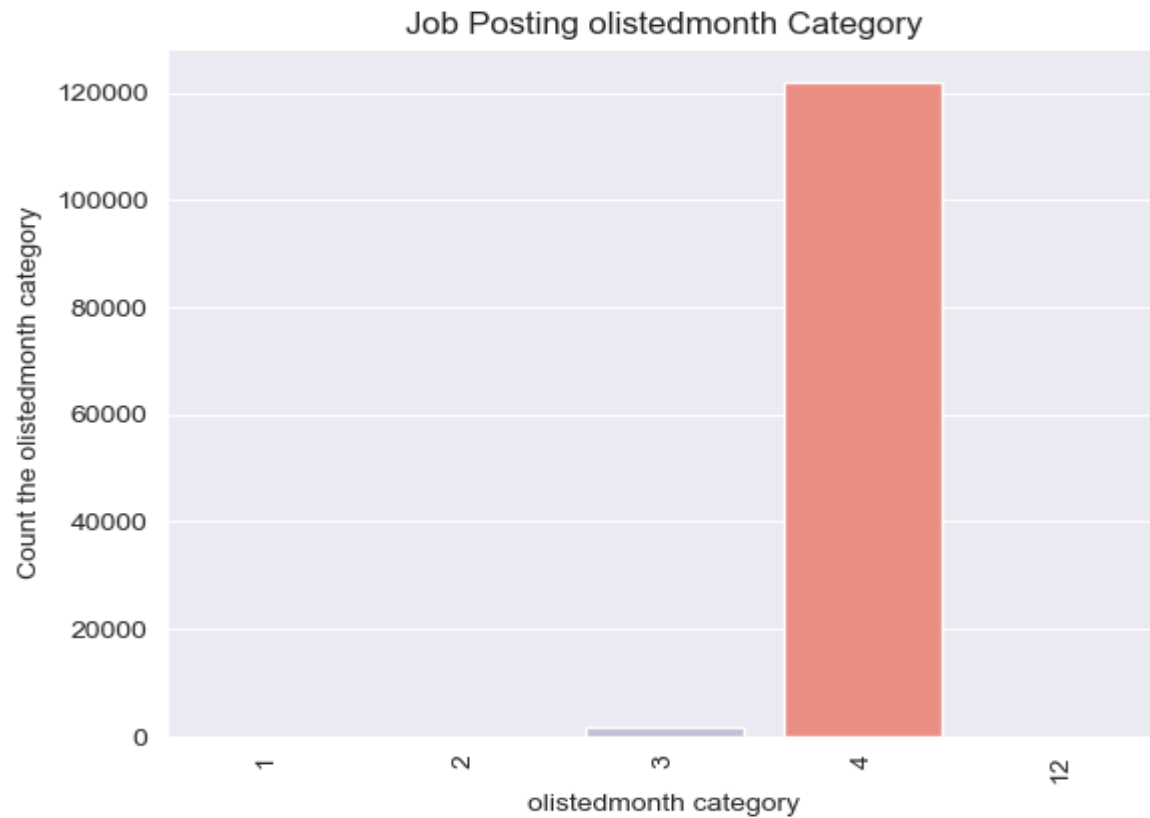
April Month is frequntly accured in the olistedmonth category.
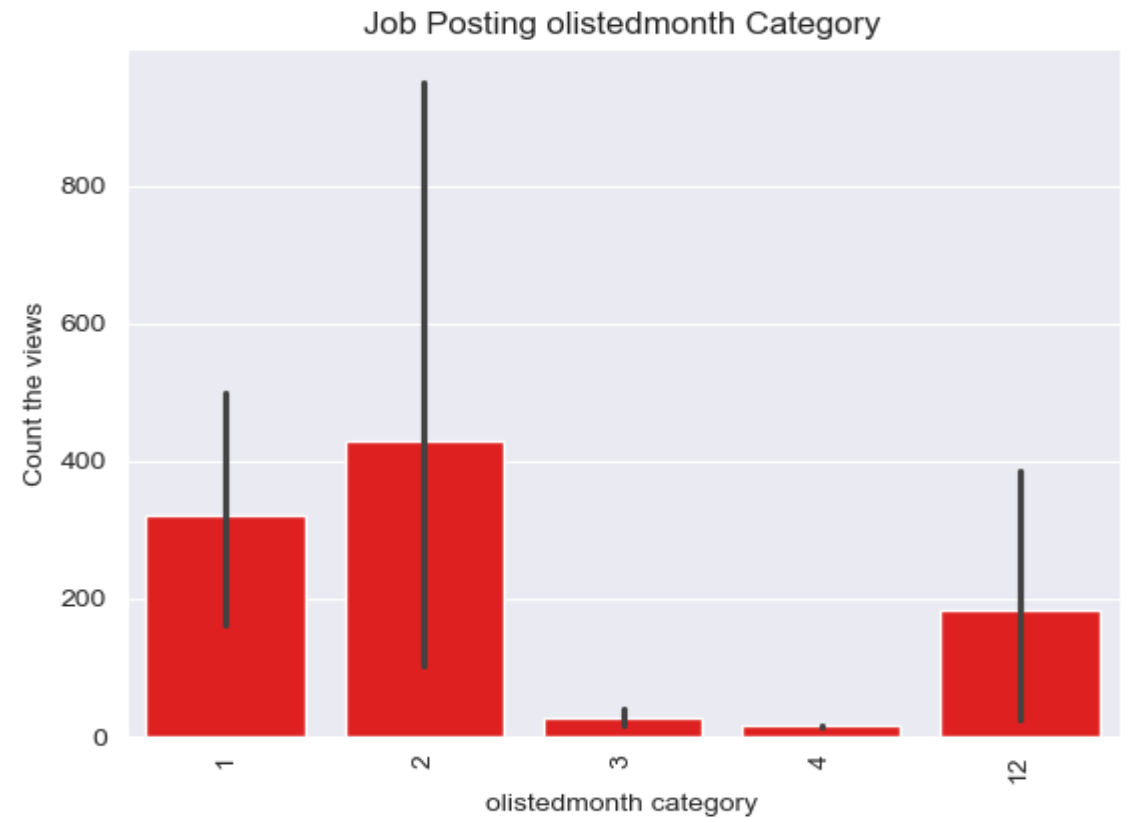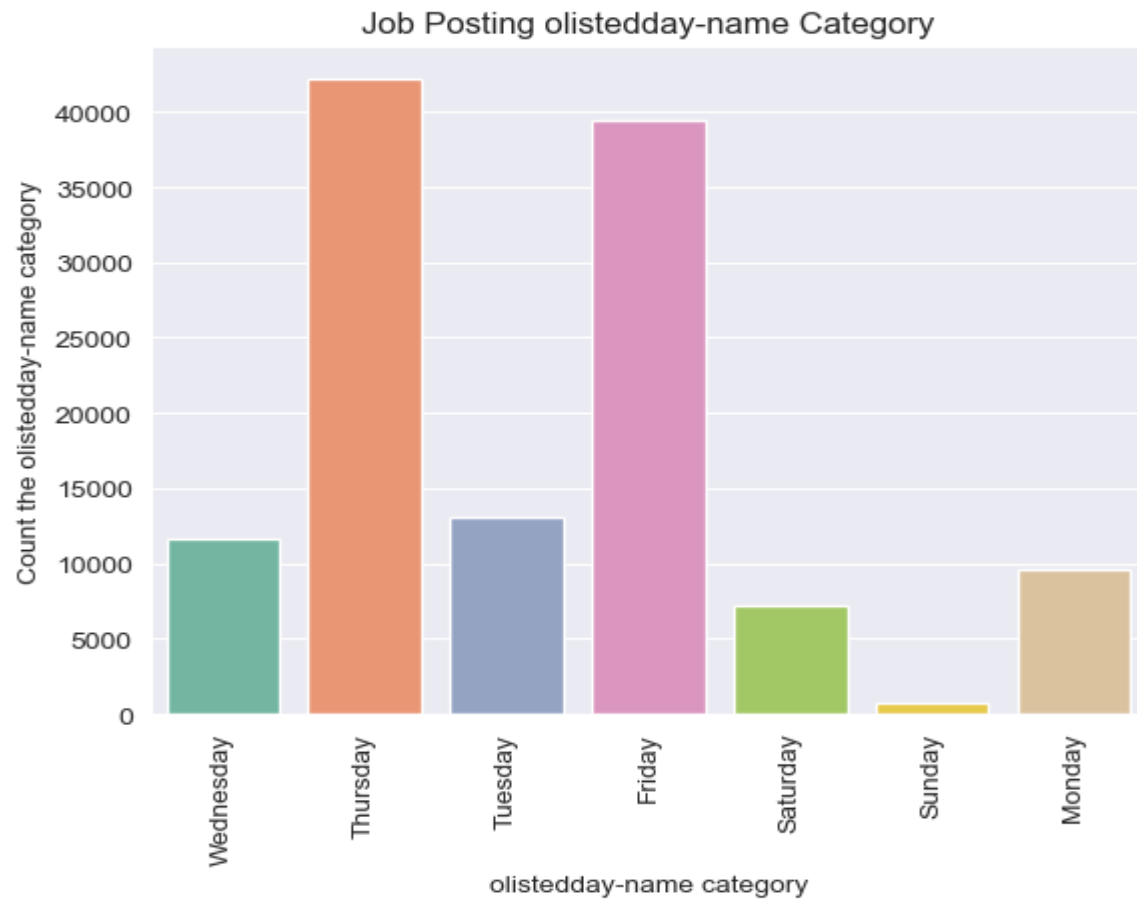
As you can see clearly in the chart in the listed month 1,12 getting the highest views but if you get more variance then go with feb month.
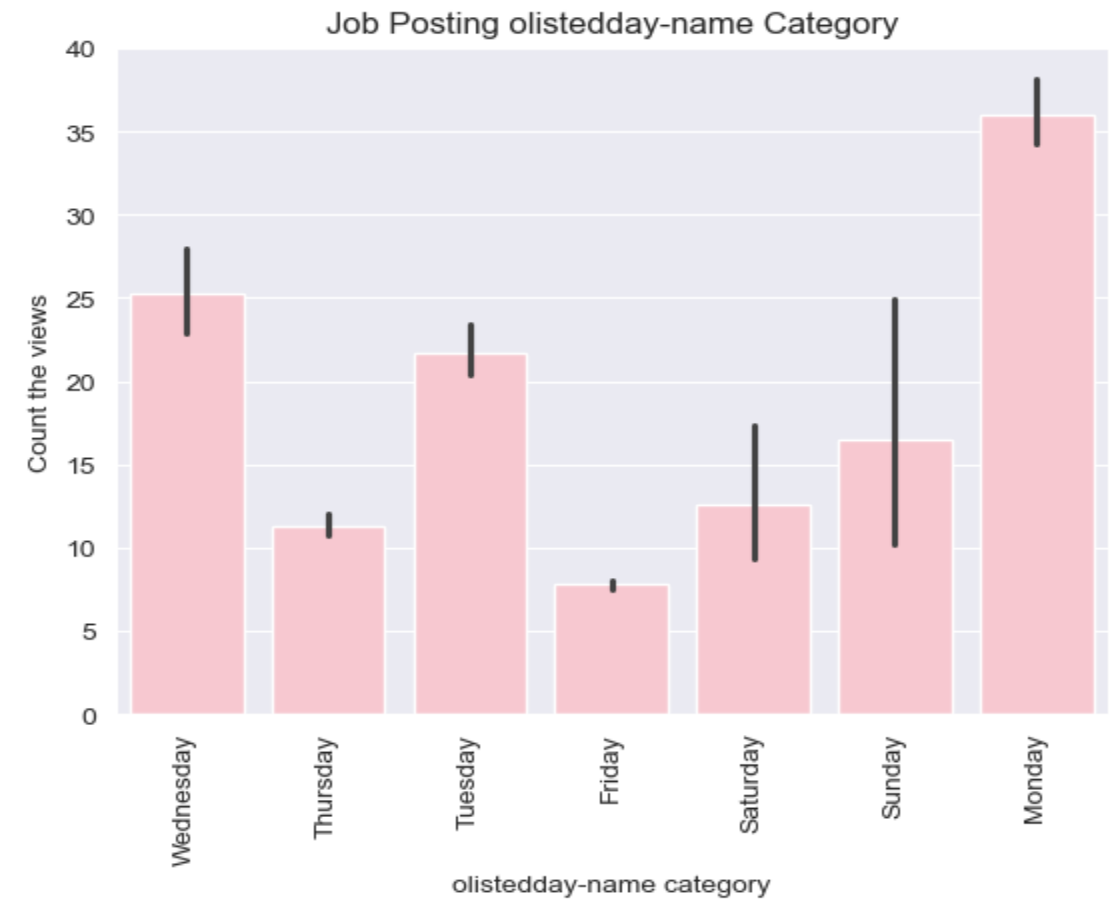
Job Posting olistedday Category

Date 18,19 are more accured in the olistedday category.
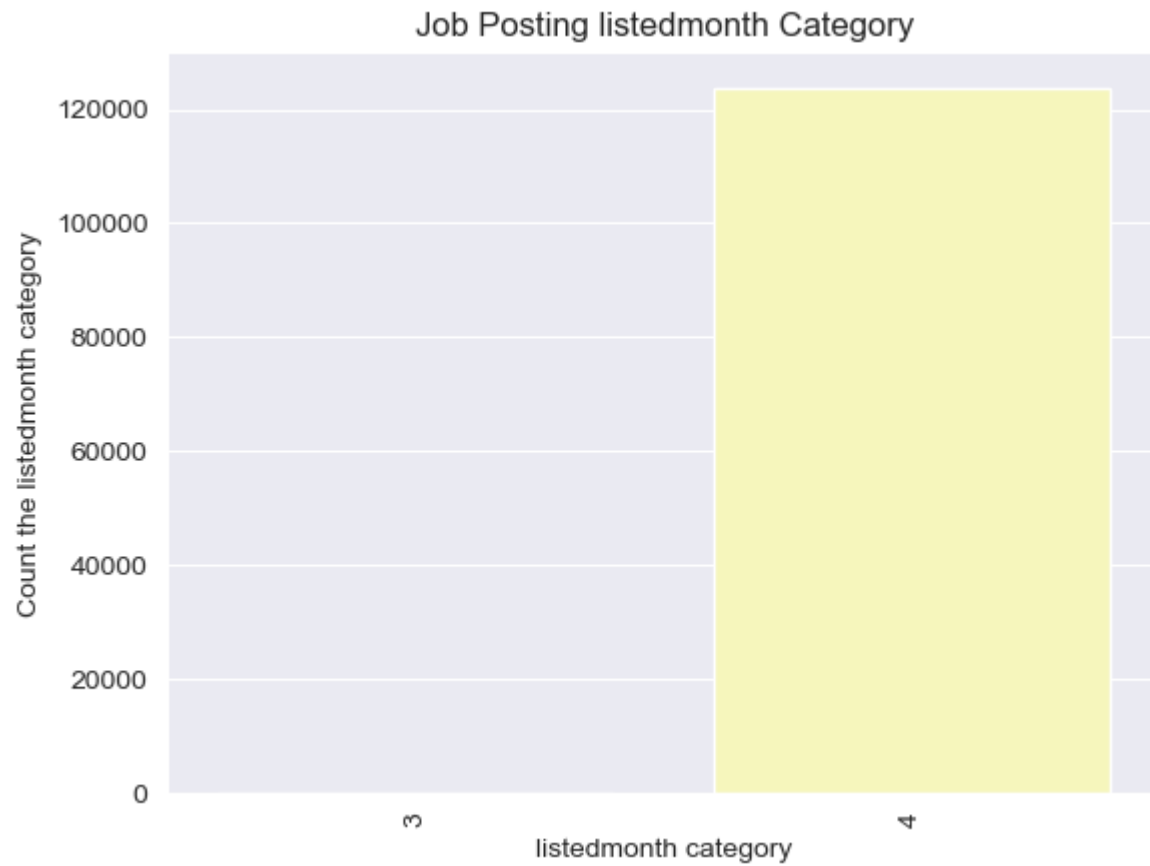
You can see clearly in the listedday category in job postings in the date of 22,23,31 getting the highest variance in views.
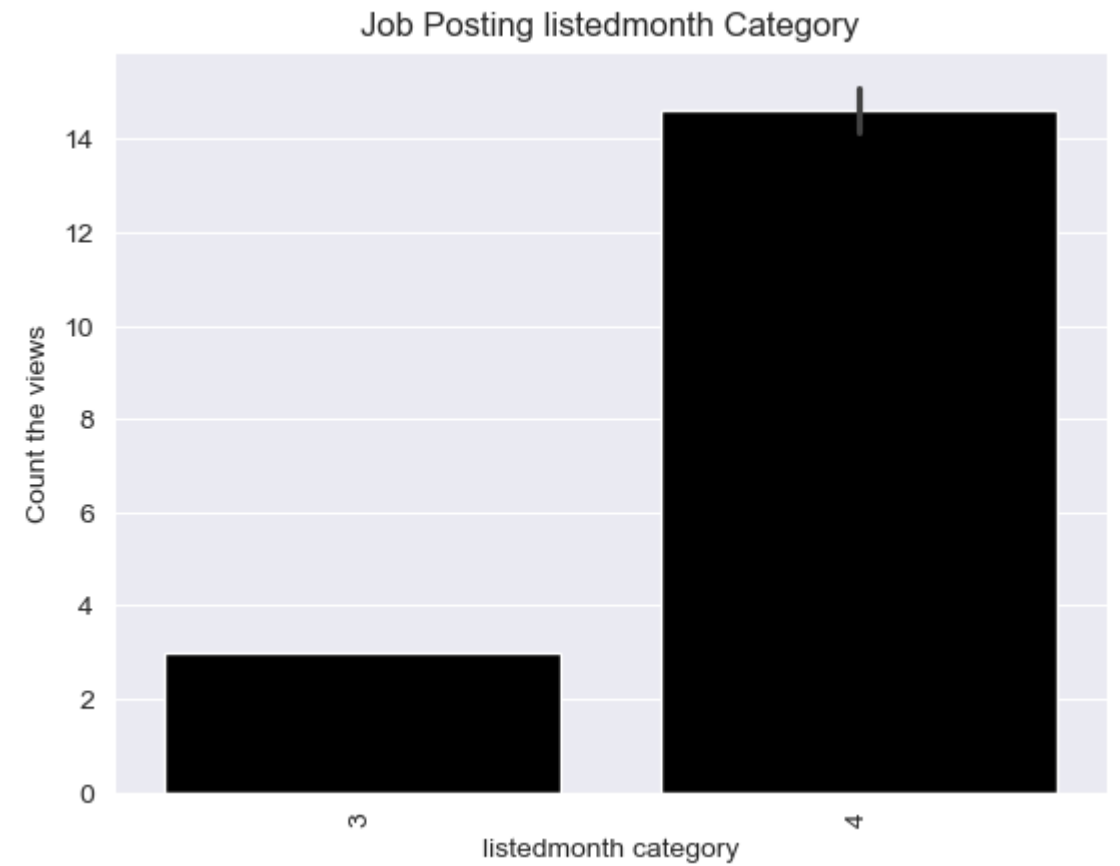
Job Posting olistedday-name Category

Thirsday is high in 0listedday-name category.

Those company's job posting listed on the day of 'Monday' getting the highest views on linkedin platform.

Job Posting listedmonth Category

Job Posting listedmonth Category

In the Listedmonth Category April month is most frequent.

You can see clearly the winner is job posting listed month is April getting the highest public views on linkedin platform.
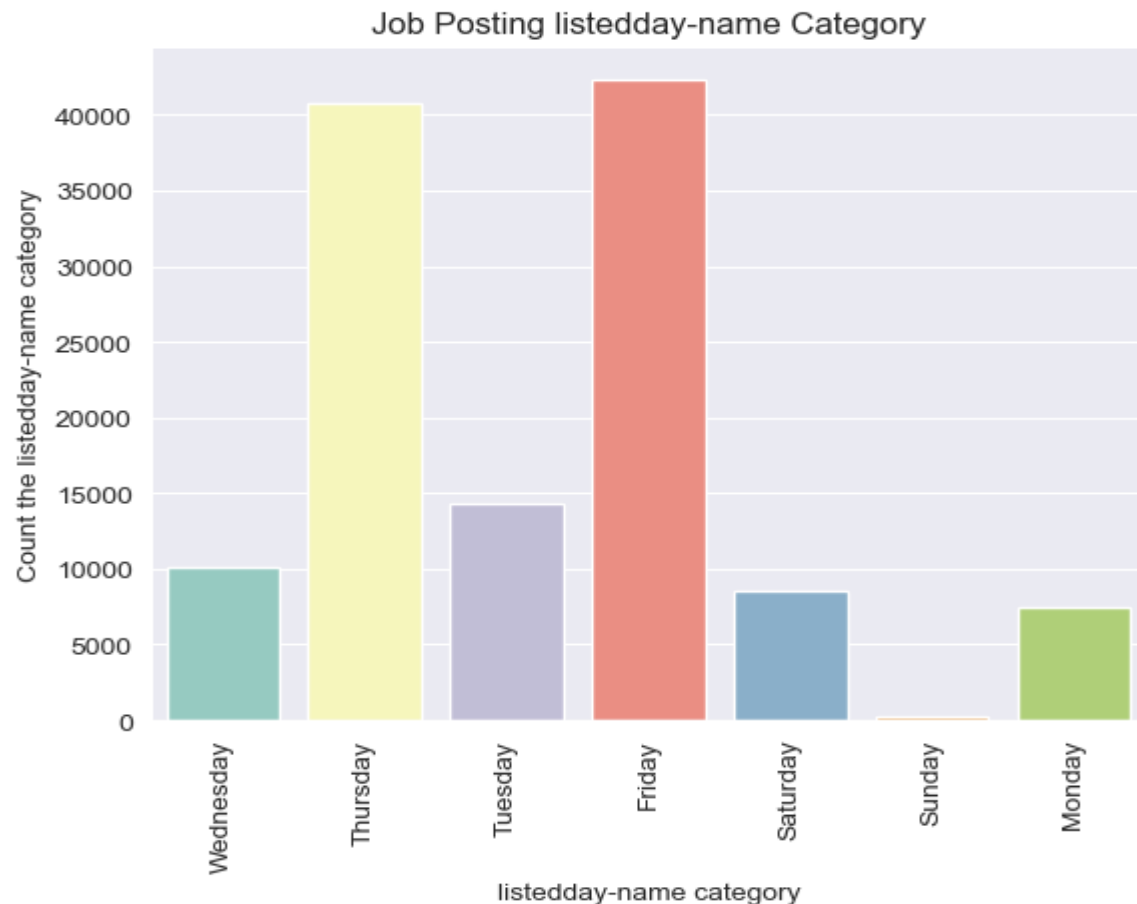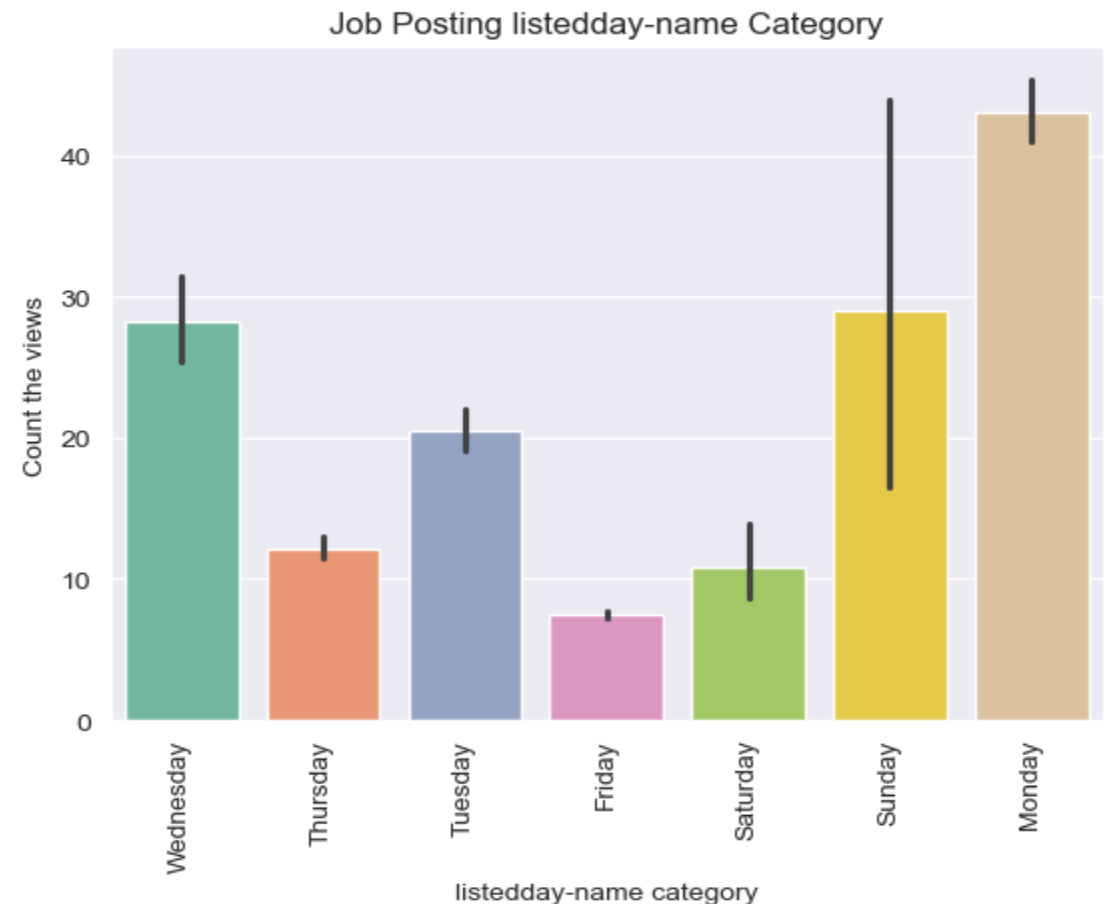
Job Posting listedday Category

In the Listedday Category 18 & 19 date is most frequent.

Those company job posting listed day in the date of 15th getting the highest views on linkedin platform.

Job Posting listedday-name Category
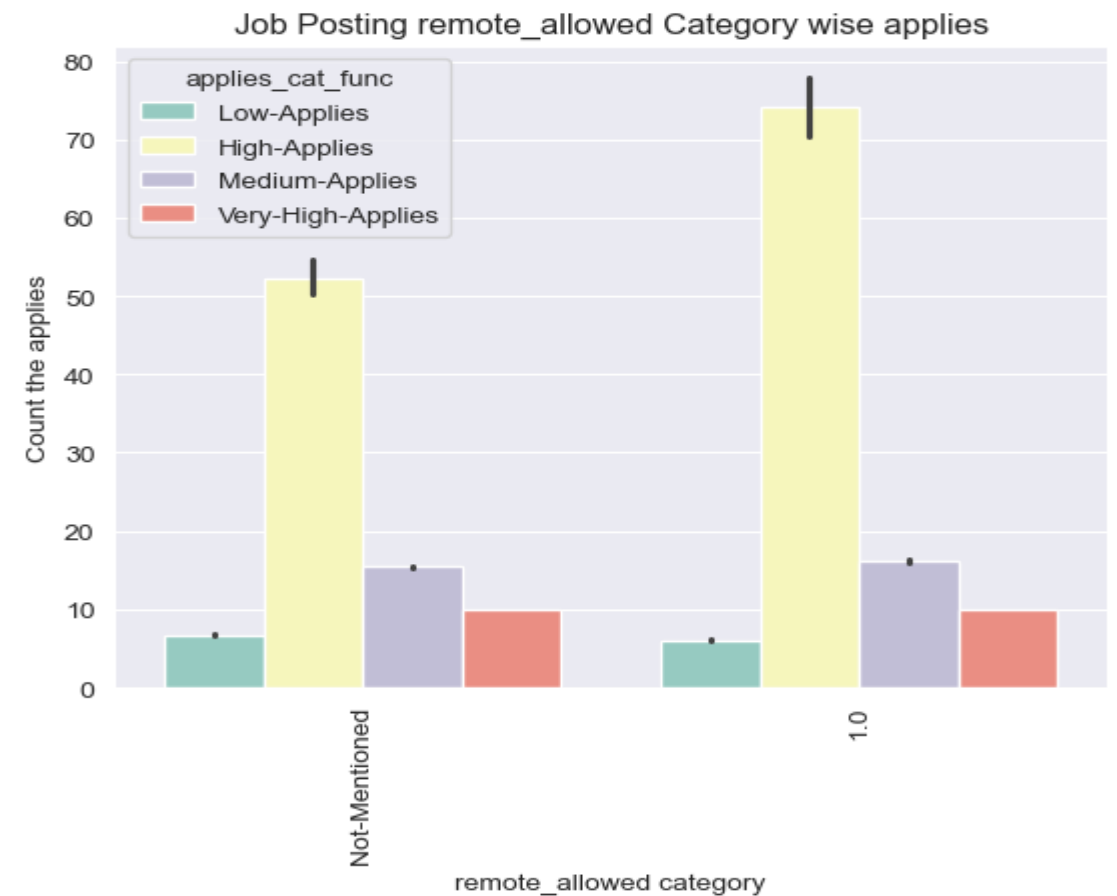
Job Posting listedday-name Category

In the Listedday-name Category Friday day is most frequent.

You can see clearly those company's job postings on the day of 'Monday' getting the highest views but "Sunday" get more variance in the views on linkedin platform.
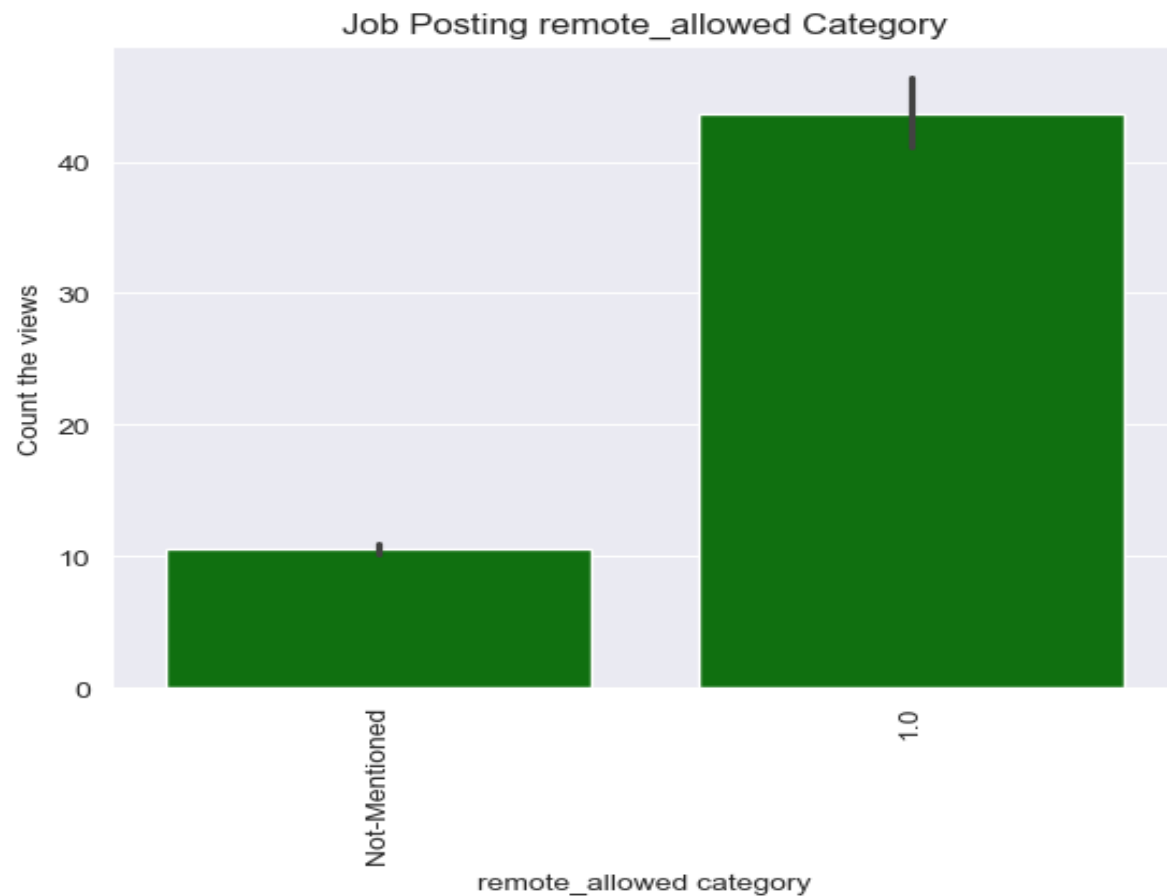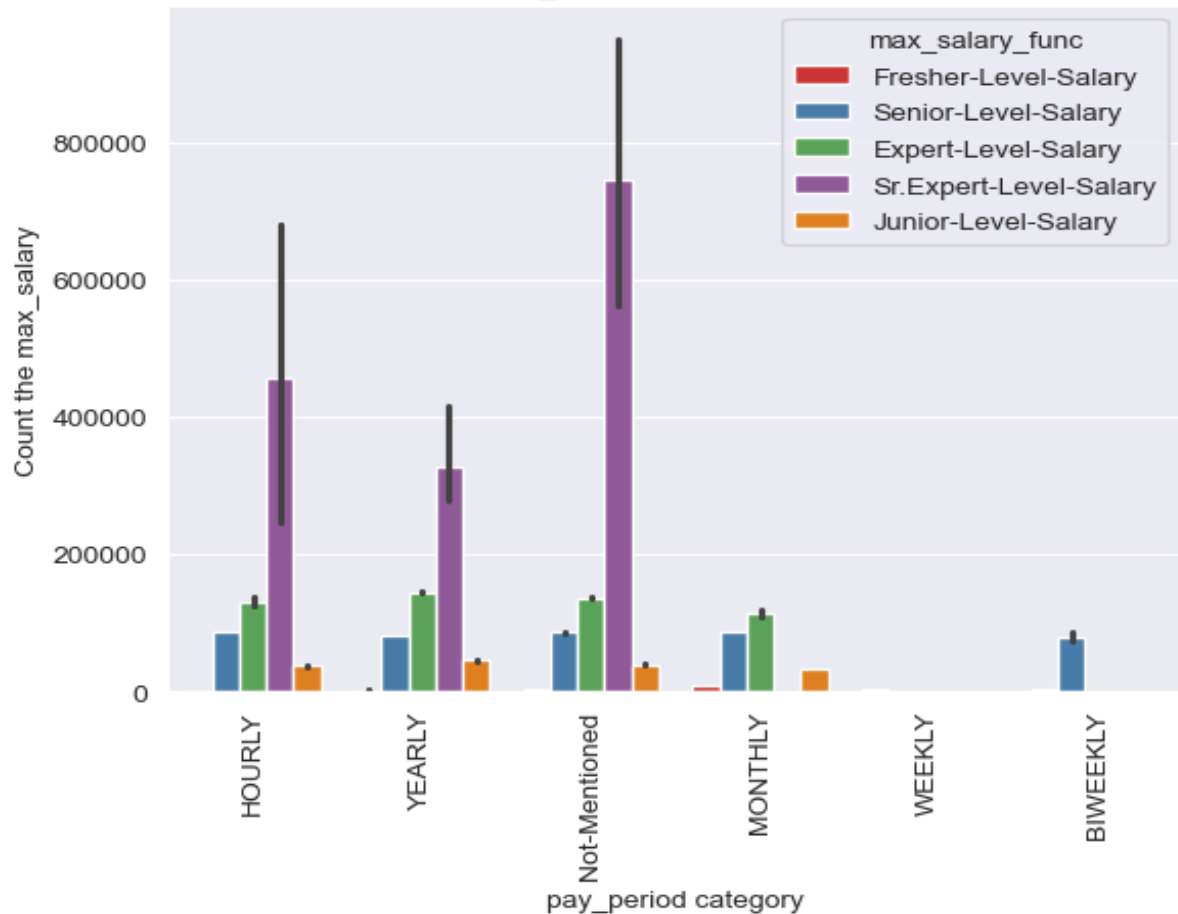
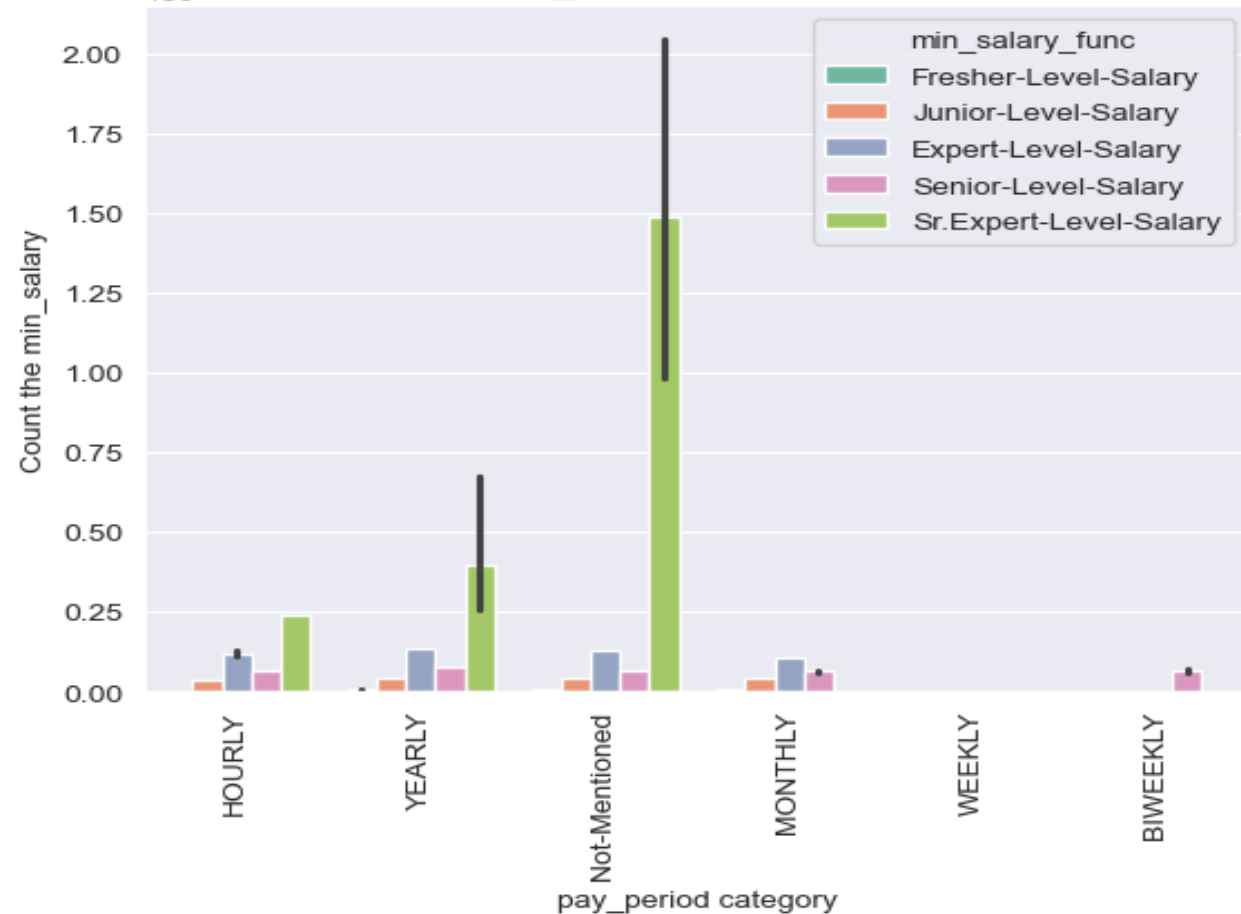**Most of the company offered remote allowed getting the highes views.**

You can see the those company offeres the remotely job allowed getting the highest applies on it.

Job Posting pay_period Category wise max salary

Job Posting pay_period Category wise min salary

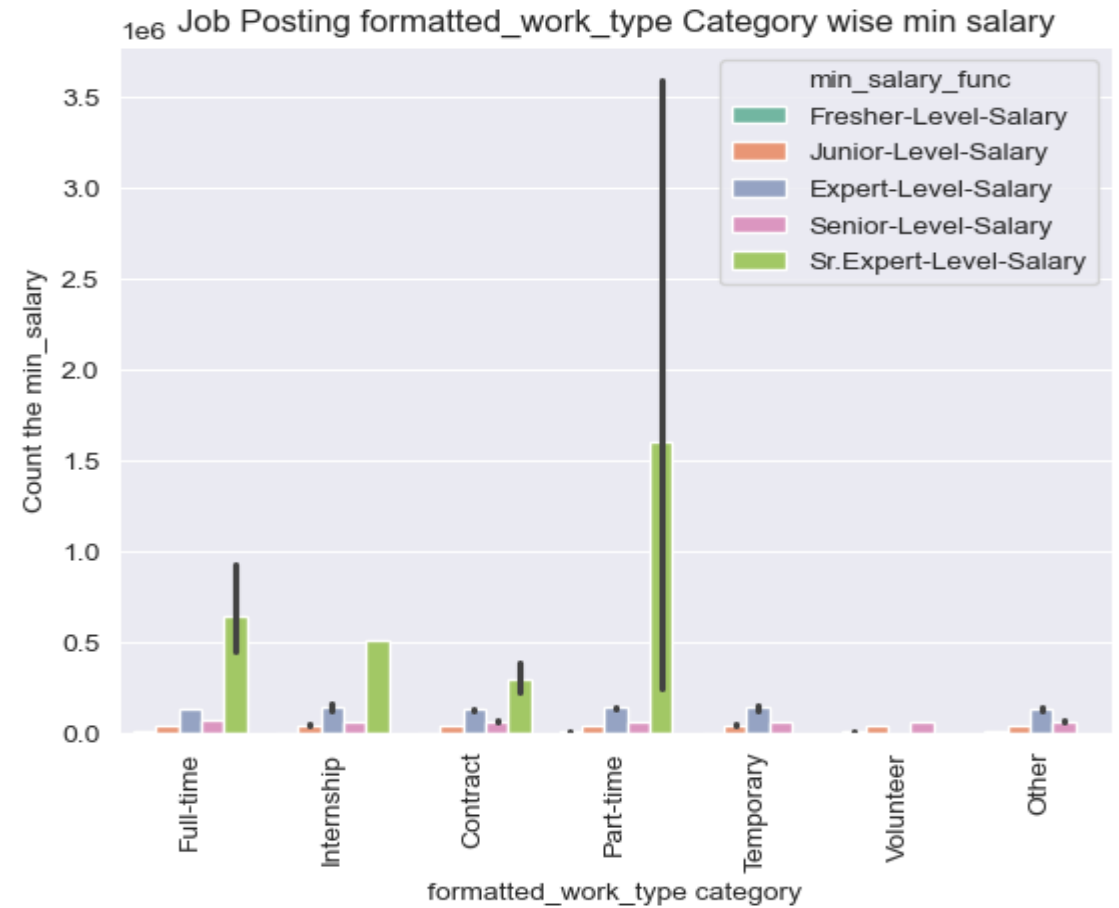**You can see sr.Expert Level Salary category getting the highest maximum salary on the 'Hourly' basis.**

Interesting thing is in the opposite side 'Sr.Expert Level' getting the highest salary in the the minimum salary category but their paytype is not-mentioned.

Job Posting formatted_work_type Category wise max salary

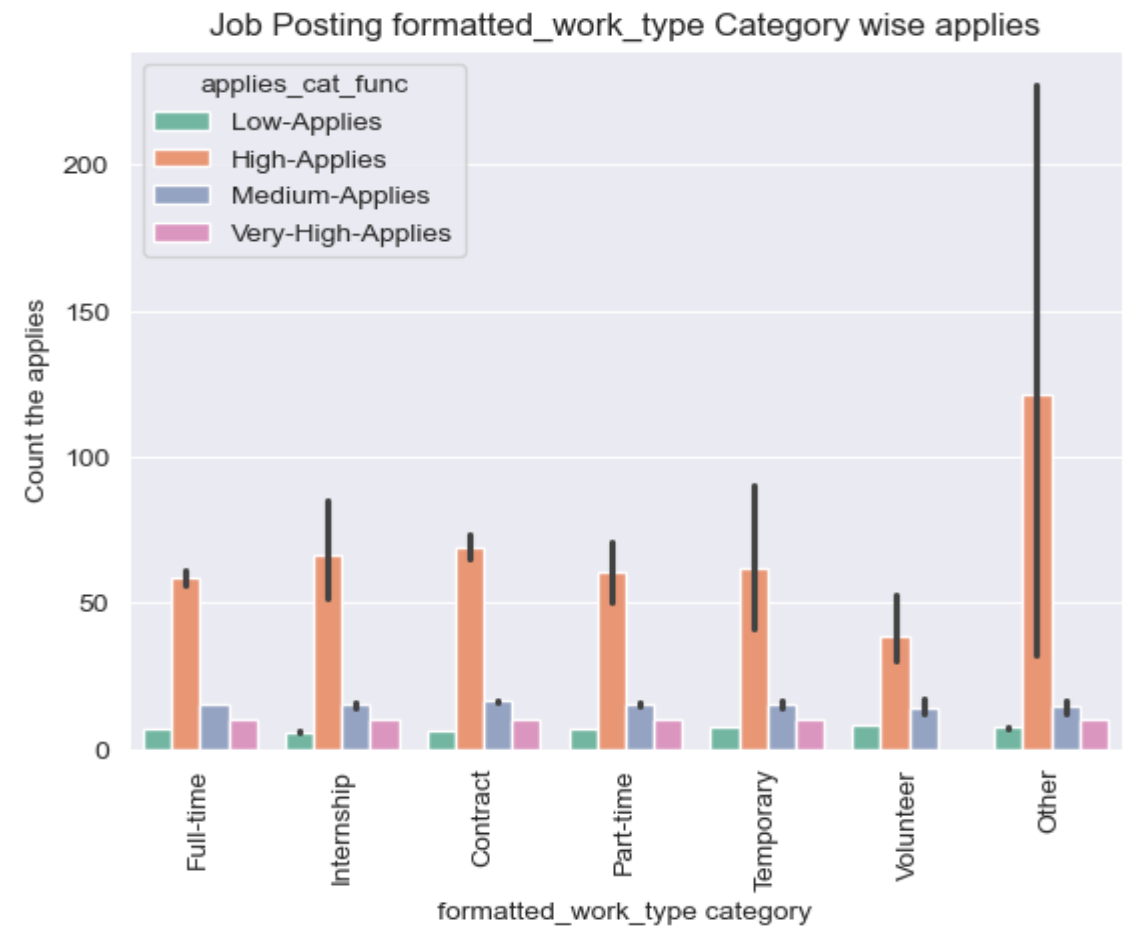Job Posting formatted_work_type Category wise min salary

As you know that 'Sr.Expert Level Salary' getting the highest max salary in all formatted work types.
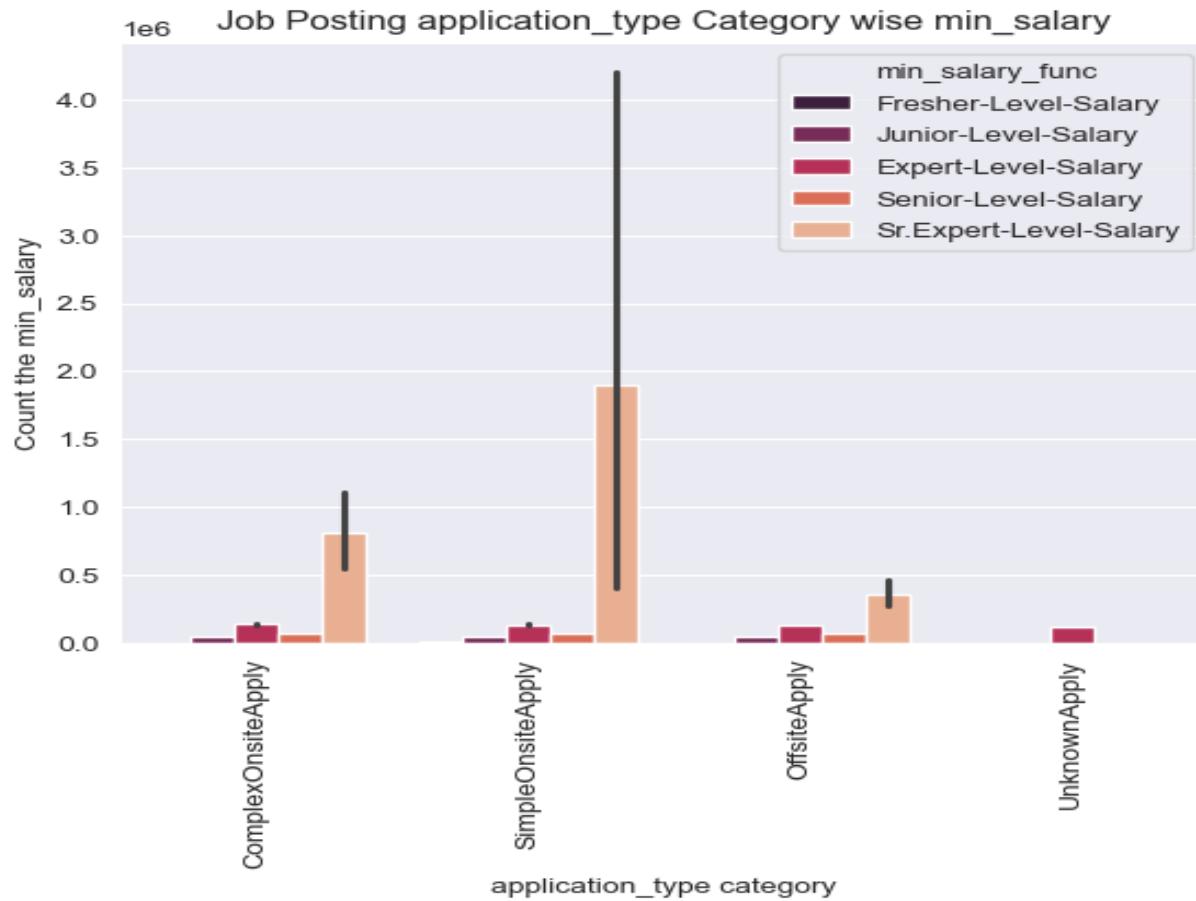
Here again the winner is 'Sr.Expert Level' category earns higher minimum salary among all the formatted work type categories.

Job Posting pay_period Category wise applies

Job Posting formatted_work_type Category wise applies
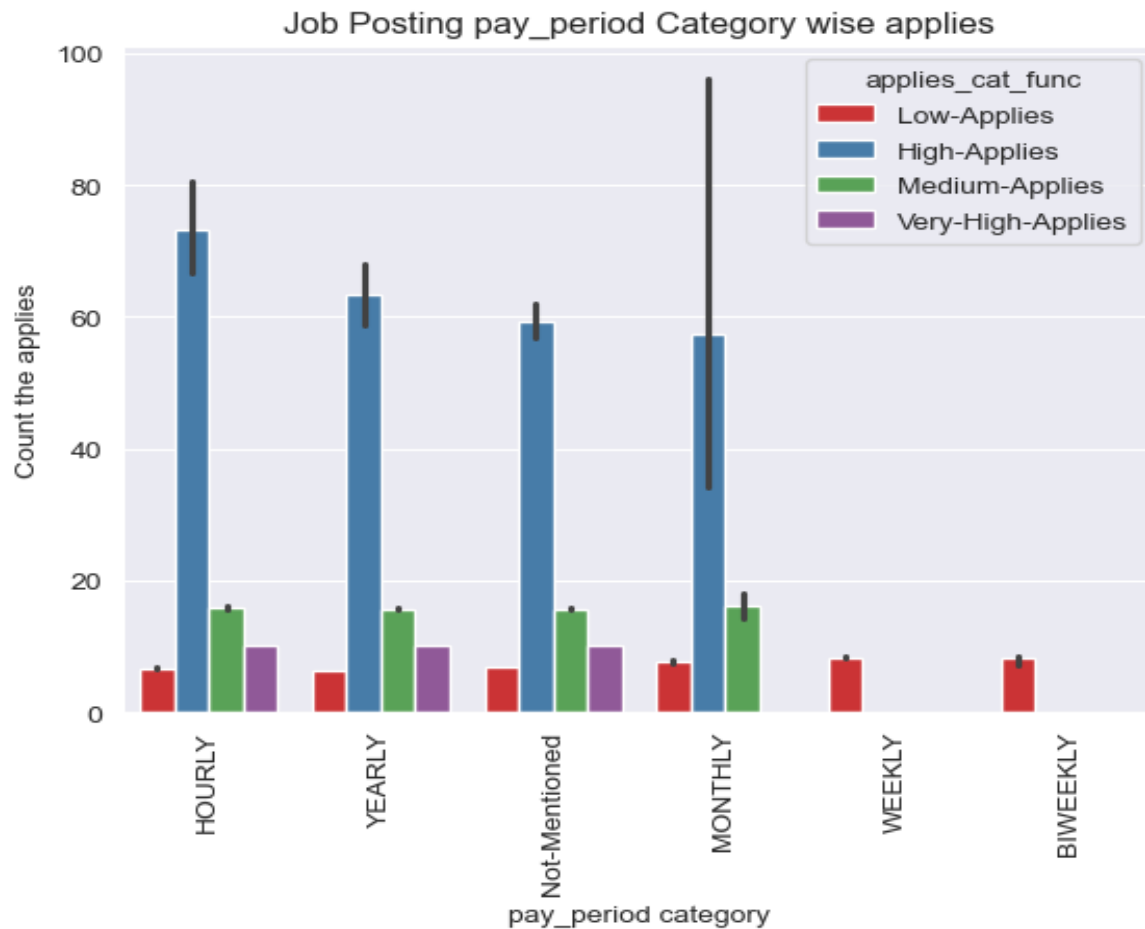
You can seel clearly company offeres salary on the basis of 'Hourly','Yearly','Not-Mentioned','Monthly' pay-period category getting the highest applies & 'Weekly','Biweekly' getting the low applies on linkedin platform.

You can see clearly those company offer the other formatted work type category getting the highest applies on it.

You can see In the application type category 'SimpleOnsiteApply' category and specially 'Sr.Expert Level' getting the highest minimum salary in this category among the all.
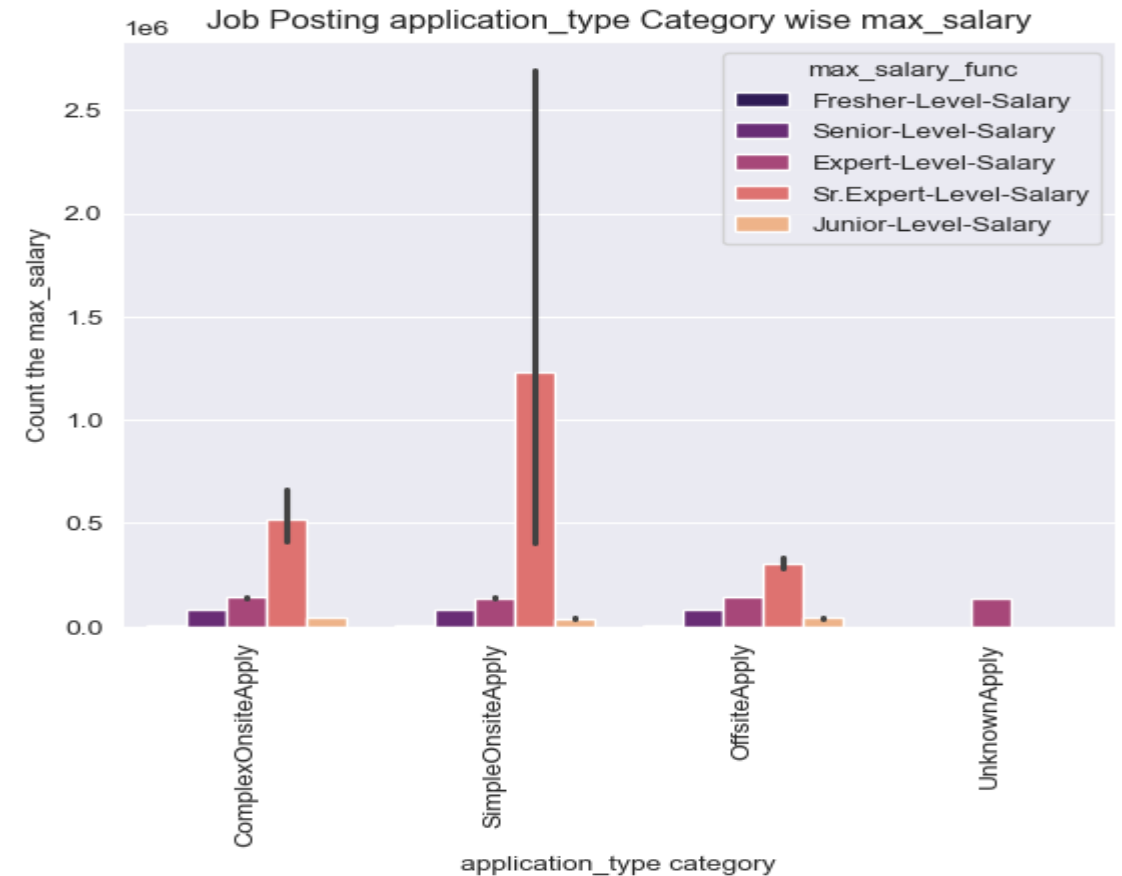
Again the winner is 'SimpleOnsiteApply' category earns more especially for the 'Sr.Expert Level' salary category.

Job Posting work_type Category wise max_salary

Job Posting work_type Category wise min_salary
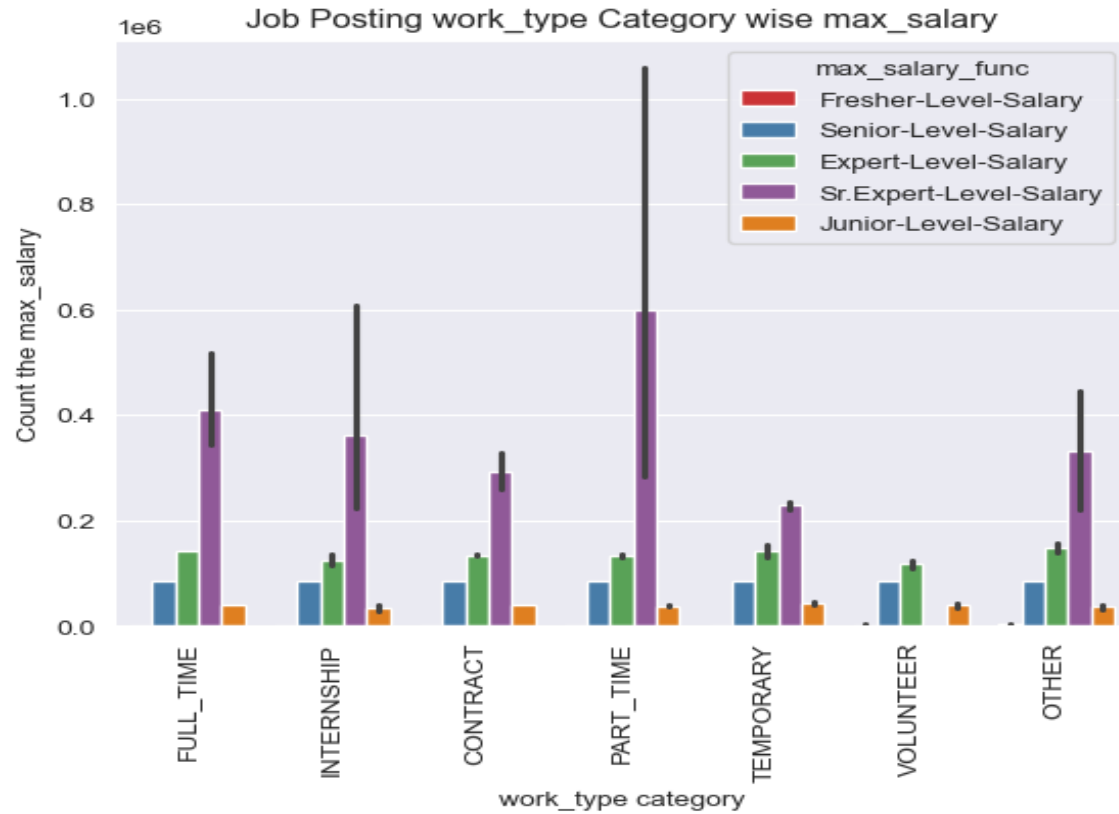
Those Work types is based on the 'Part_Time' category and in the 'Sr.Expert Level Salary' earns more maximum salary rest of all.
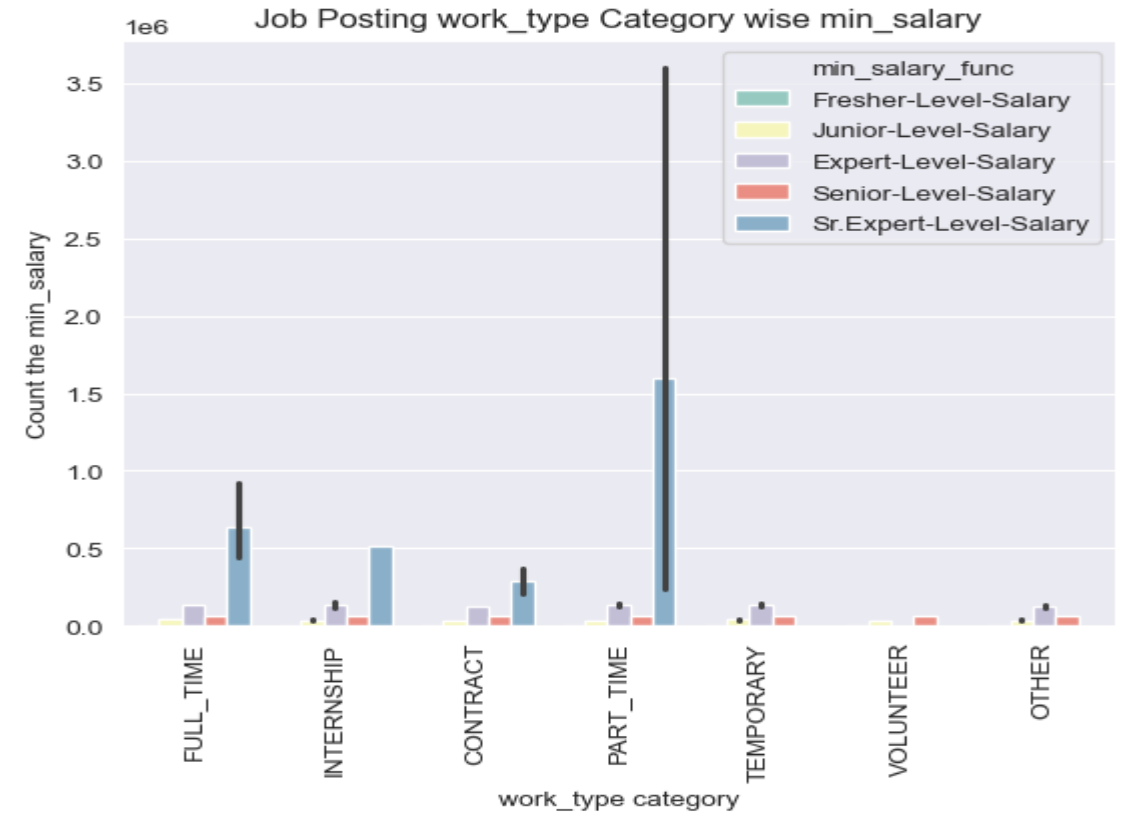
Even the minimum salary is high for 'PART_TIME' jobers but only for 'Senior Expert Level' Salary.

Common Insights and Recommendations

1) Focus on Sales Manager roles and highlight Senior Expert Level salaries.

2) Prioritize Full-Time and Remote-Allowed jobs to attract maximum applications.

3) Mention Base Salary clearly to improve candidate interest and trust.

4) Post jobs on Mondays, especially around dates like 15th–19th, to get higher visibility.

5) Offer Yearly or Monthly salary pay periods instead of Weekly or Biweekly.

6) Prefer OffsiteApply and ComplexOnsiteApply methods for job applications.

6) Prefer OffsiteApply and ComplexOnsiteApply methods for job applications.

7) Post jobs more in January, February, December, and April for better reach.

8) Senior Expert Level jobs earn the highest minimum and maximum salaries.

9) Jobs with Not-Mentioned Compensation still perform well, but Base Salary detail attracts better candidates.

10) Part-Time jobs with Senior Expert Level offer the highest salaries in that category.

11) Use popular title keywords like Manager, Engineer, Sales to match candidate searches.

12) Offering Remote Jobs significantly increases application rates.

Thank You