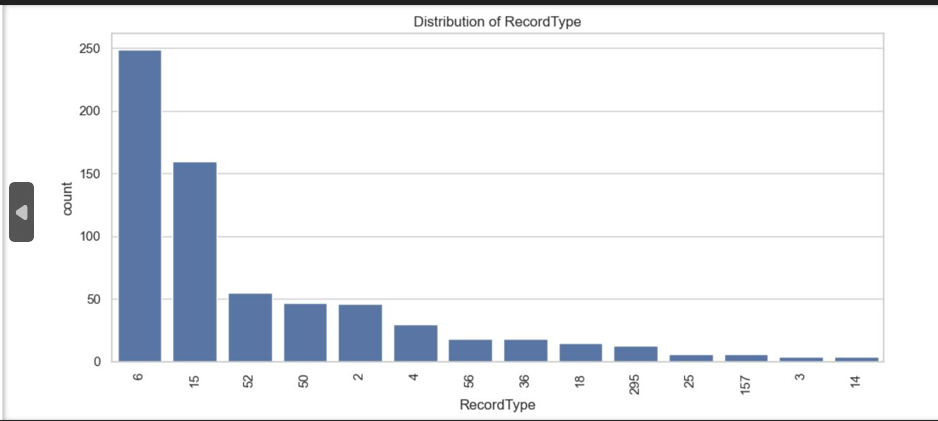
**Univariate Analysis Report** based on the bar charts, focusing on categorical columns from dataset:

1. **Distribution of RecordType :**



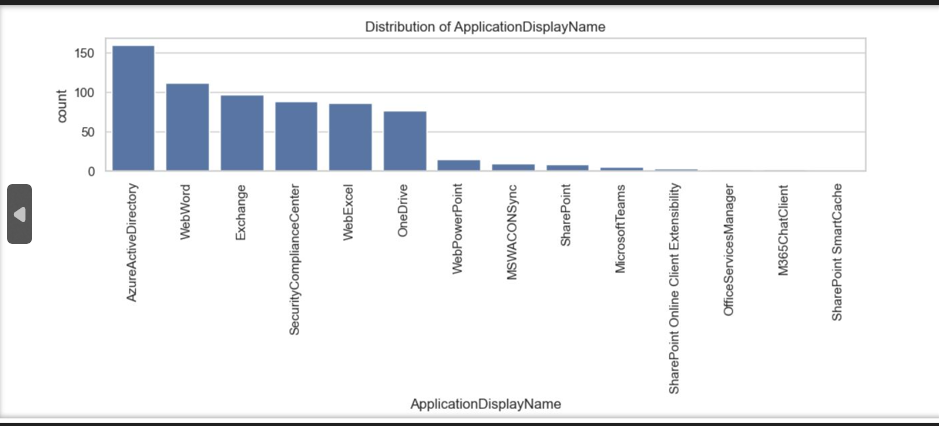
**Chart Summary:**

* Most frequent record type is 6 with a count of ~250.
* Followed by 15 (~160) and then a steep drop to others like 52, 50, 2, etc.
* Long tail includes many less frequent types (e.g., 295, 25, 3, 14).

**Insights:**

* Record Type 6 and 15 dominate the dataset.
* Suggests majority of events/logs pertain to these two types.
* This skew could impact model training if RecordType is a feature—consider balancing if needed.

1. **Distribution of ApplicationDisplayName**



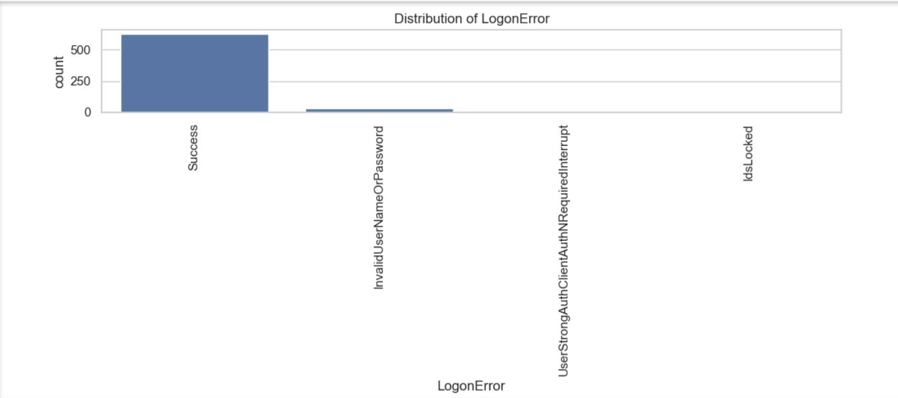
**Chart Summary:**

* Top 5 applications:
  1. AzureActiveDirectory (~155)
  2. WebWord (~110)
  3. Exchange
  4. SecurityComplianceCenter
  5. WebExcel
* Applications like SharePoint SmartCache, OfficeServicesManager, M365ChatClient, etc., have very low occurrences.

**Insights:**

* Microsoft 365 services dominate the dataset, especially identity and office apps.
* **AzureActiveDirectory** logs could indicate a focus on authentication or access-related events.
* Less used apps might not be significant but could be outliers worth monitoring.

**3. Distribution of LogonError**



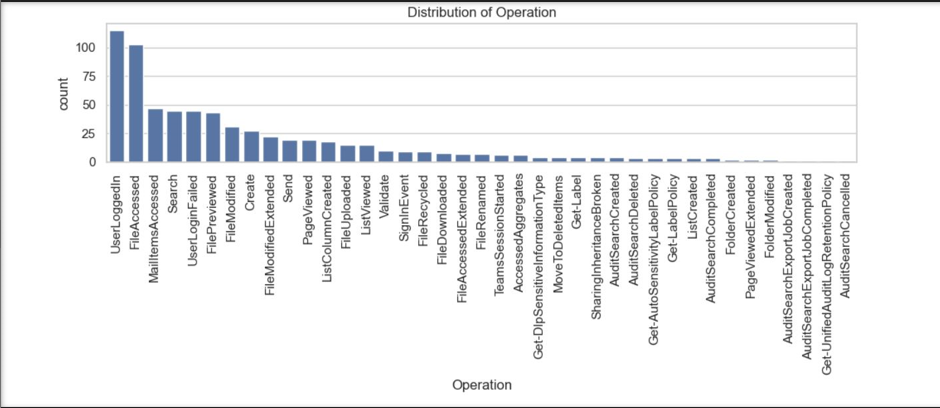
**Chart Summary:**

* Vast majority are Success (~550+).
* Very few error types present:
  + InvalidUserNameOrPassword
  + UserStrongAuthClientAuthNRequiredInterrupt
  + IdsLocked

**Insights:**

* Successful logons dominate the logs.
* While error rates are low, error types like IdsLocked and InvalidUserNameOrPassword are important **security signals**.
* Important for threat detection models—errors should be monitored and possibly over-sampled for model training.

**4. Distribution of Operation**



**Chart Summary:**

* Most common operations:
  + UserLoggedIn (~110)
  + FileAccessed (~100)
  + MailItemsAccessed, Search, UserLoginFailed
* Long tail includes many rare operations like:
  + AuditSearchExportJobCompleted, AuditSearchCancelled, Get-UnifiedAuditLogRetentionPolicy, etc.

**Insights:**

* Top actions are authentication (UserLoggedIn, UserLoginFailed) and content interaction (FileAccessed, MailItemsAccessed).
* Several low-frequency operations likely tied to admin/config tasks.
* Could consider grouping rare operations under “Others” for analysis/modeling.

**5. Distribution of ResultStatus**



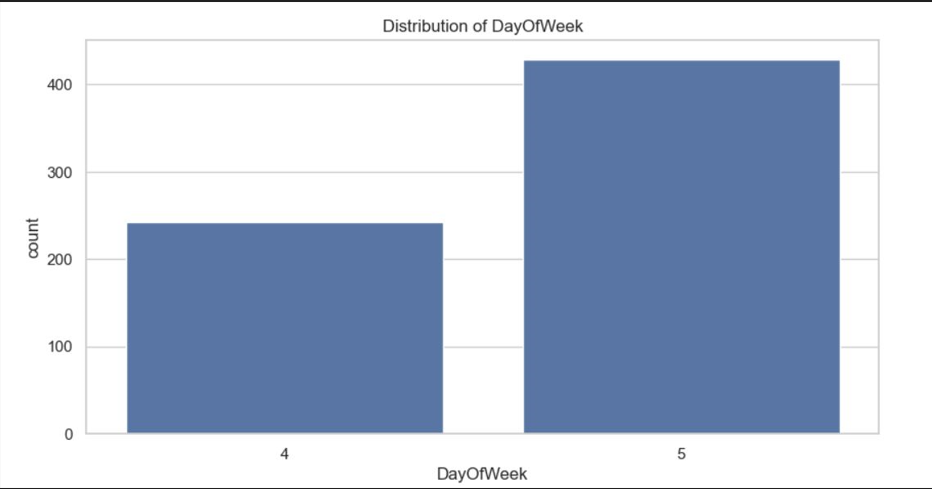
**Chart Summary:**

* succeeded: ~540
* unknown: ~130

**Insights:**

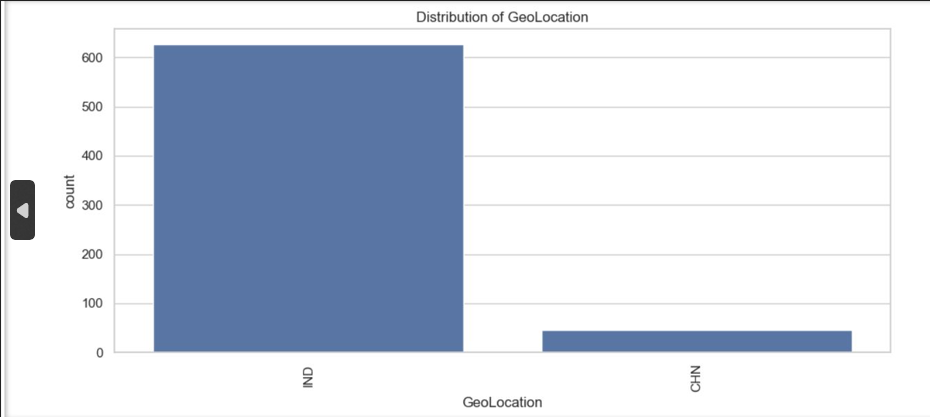
* Most operations are successfully completed.
* unknown result status might imply:
  + Incomplete events
  + Failures or untracked outcomes
* Worth investigating further to understand context or fix potential logging gaps.

**1. DayOfWeek Distribution**



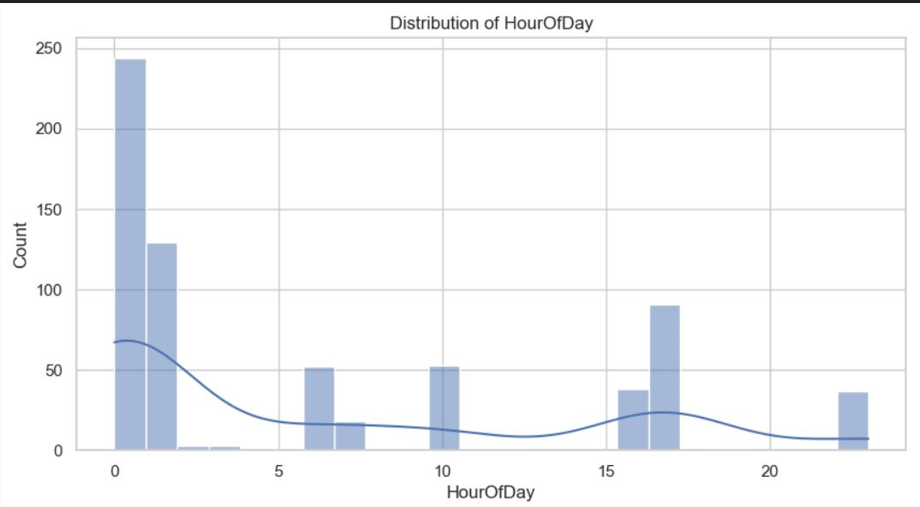
* **Analysis:**
* The distribution shows activity counts across days of the week
* Days 4 and 5 (likely Thursday and Friday) show significantly higher counts than other days
* The count drops sharply on other days (0-3, likely Sunday-Wednesday)
* Possible interpretation: Higher activity occurs later in the work week

1. **GeoLocation Distribution**



* **Analysis:**
* The variable appears to be binary with two categories: "G" and "N"
* "G" location has significantly more counts (approximately 550) compared to "N" (approximately 50)
* Imbalanced distribution suggests "G" is the dominant or default location
* Possible interpretation: "G" might represent a headquarters or primary office location

1. **HourOfDay Distribution**



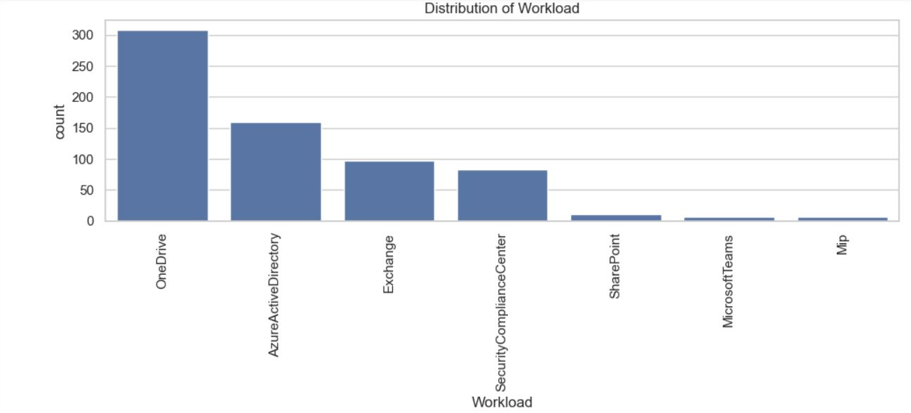
* **Analysis:**
* Shows activity distribution across 24 hours
* Peak activity occurs around hour 10 (10 AM)
* Secondary peak around hour 15 (3 PM)
* Minimal activity during nighttime hours (0-5, midnight to 5 AM)
* Typical workday pattern with morning and afternoon peaks

1. **IsWeekend Distribution**



* **Analysis:**
* Compares activity on weekdays (False) vs weekends (True)
* Weekday activity is dramatically higher (approximately 375 counts) than weekend activity (approximately 25 counts)
* Ratio of about 15:1 (weekday:weekend)
* Clear business/work pattern with minimal weekend activity

1. **Workload Distribution**



* **Analysis:**
* Shows distribution across different workload types
* SharePoint and Exchange follow with substantial counts (~200 and ~175 respectively)
* OneDrive, AzureActiveDirectory, and SecurityComplianceCenter have moderate counts (~100-150)
* "Mp" has the lowest count (~50)
* Possible interpretation: Teams is the most heavily used service in this environment