

**MALAYALAM FREQUENCY COUNT**  
Study Report (Character Frequency)

By

S.Prama, Manu Joseph  
Department of Linguistic University of Kerala

**മലയാള ആവൃത്തി കണ്ടെത്തൽ**  
പഠന റിപ്പോർട്ട് (ക്യാരക്ടർ ആവൃത്തി)

By

എസ് പ്രേമ  
മനു ജോസഫ്

ഭാഷാ ശാസ്ത്ര വിഭാഗം കേരള സർവകലാശാല

# 1. ആമുഖം

മലയാളഭാഷാസാങ്കേതികവിദ്യ അതിന്റെ ശൈശവാവസ്ഥയിലാണ്. അടിസ്ഥാനപരമായി ആവശ്യമായ ഭാഷാഗവേഷണം നടത്താതെയുള്ള നിർമ്മാണപ്രവർത്തനങ്ങളാണ് ഇന്ന് നടക്കുന്നത്. ഭാഷാസാങ്കേതികവിദ്യാഗവേഷണത്തിന്റെ ആദ്യപടി ഭാഷയുടെ ഭൗതികസത്ത( material substance)യെപ്പറ്റിയുള്ള അന്വേഷണമാണ്. ഈ അന്വേഷണത്തിന് രണ്ട് തലങ്ങളുണ്ട്. ഒന്ന്, ഭാഷയുടെ ശബ്ദസത്ത(phonetic substance)യെപ്പറ്റിയുള്ള അന്വേഷണം, രണ്ട്, എഴുത്തിന്റെ ദൃശ്യതല (graphic substance)ത്തെപ്പറ്റിയുള്ള അന്വേഷണം. ലിഖിതതലത്തിലെ ഭാഷയിൽ ഉപയോഗിക്കുന്ന അക്ഷരങ്ങളുടെ ആവൃത്തിനിർണ്ണയമാണ് ഈ പദ്ധതിയുടെ ഉദ്ദേശ്യം. ഇത് ഒരു തുടക്കം മാത്രമേ ആകുന്നുള്ളൂ. എല്ലാത്തരം ഗവേഷണസംരംഭങ്ങളും നിർമ്മാണത്തെ നേരിട്ടുസഹായിരിക്കുന്നതായിരിക്കണം എന്ന വാദം 'ചോറുണ്ണുവാൻ വിത്ത് വിതക്കേണ്ടതില്ല, അരി പുറത്ത് നിന്ന് വാങ്ങിയാൽ മതി എന്ന് പറയുന്നതുപോലെയാണ്'.

ഈ പദ്ധതിയുടെ കണ്ടെത്തലുകൾക്ക് താഴെ പറയുന്ന ഉപയോഗസാധ്യതകളുണ്ട്.

1. മലയാളഭാഷാസാങ്കേതികവിദ്യയിൽ നാമിന് വളരെ അഭിമുഖീകരിക്കുന്ന പ്രശ്നം മലയാളം കീബോർഡിന്റെ പ്രശ്നമാണ്. ആൽഫബെറ്റിക് എഴുത്തുവ്യവസ്ഥയ്ക്കു രൂപപ്പെടുത്തിയ കമ്പ്യൂട്ടർ കീബോർഡ് മലയാളഭാഷയുടെ ലിപികളെ ഉൾക്കൊള്ളുവാനോ മലയാളഭാഷാലിപി സങ്കീർണ്ണതയെ ഈ കീബോർഡിലേക്ക് ചുരുക്കുവാനോ കഴിയുന്നതല്ല. അങ്ങനെ ചുരുക്കി ഉപയോഗിക്കുന്നത് കൊണ്ടാണ് മലയാളം ടൈപ്പിംഗ് ബുദ്ധിമുട്ടുള്ള ജോലിയായി സാധാരണ കമ്പ്യൂട്ടർ ഉപയോക്താക്കൾക്ക് അനുഭവപ്പെടുന്നത്. ഈ പ്രശ്നം താഴെ പറയുന്ന മാർഗങ്ങൾ വഴി പരിഹരിക്കാവുന്നതാണ്.
2. കീബോർഡ് പരിഷ്കരിക്കുക.
3. കീബോർഡ് ഒഴിവാക്കി ശബ്ദത്തെ ടെക്സ്റ്റ് ആക്കി മാറ്റുന്ന സംവിധാനം നിർമ്മിക്കുക.
4. മൗസ് പ്രിന്റിംഗ് സംവിധാനം ഉപയോഗിക്കുക.
5. ഇംഗ്ലീഷ് അക്ഷരങ്ങൾ വഴിയുള്ള മലയാളം എൻട്രി സംവിധാനം ഉപയോഗിക്കുക.

1971-ൽ ടൈപ്പ് റൈറ്റർ കീബോർഡ് പരിഷ്കരിക്കാൻ നടത്തിയ ആവൃത്തിനിർണ്ണയത്തിലെ കണ്ടെത്തലുകളിൽ നിന്നും ഏറെ വ്യത്യസ്തമാണ് ഈ പഠനത്തിലെ കണ്ടെത്തലുകൾ. അതായത് മലയാള ഭാഷ ഏറെ മാറിയെന്നർത്ഥം. മാറ്റത്തിന്റെ കാരണങ്ങളെപ്പറ്റി ഏറെ അന്വേഷണങ്ങൾ നടക്കേണ്ടതുണ്ട്. ഈ കണ്ടെത്തലുകൾ കീബോർഡ് ഇനിയും പരിഷ്കരിക്കേണ്ടതിന്റെ ആവശ്യകതയെ ശ്രദ്ധയിൽ കൊണ്ടുവരുന്നു. അതായത് കൂടുതൽ ആവൃത്തി വരുന്ന അക്ഷരങ്ങൾ കൂടുതൽ അനായാസമായി ടൈപ്പ് ചെയ്യാവുന്നതരത്തിൽ കീബോർഡിൽ പുനർവിതരണം ചെയ്യേണ്ടതുണ്ട്.

മലയാളഭാഷാസാങ്കേതികവിദ്യാ വികസനത്തിന്റെ ആദ്യപടി എന്ന നിലയിൽ ഈ ഗവേഷണസംരംഭം നടത്തുന്നതിനായി സാഹചര്യമൊരുക്കിയ സി-ഡിറ്റ് ഡയറക്ടറും ഞങ്ങളുടെ പ്രിയപ്പെട്ട അധ്യാപകനുമായ ശ്രീ അച്യുതശങ്കർ. എസ്. നായർ, സി-ഡിറ്റിലെ ലിംഗിസ്റ്റിക് കമ്പ്യൂ

ട്ടിംഗ് ടീം ലീഡറും കലാകാരനുമായ രജിസ്ട്രാർ ശ്രീ അശോക്കുമാർ  
എന്നിവരോടുള്ള നന്ദിയും കടപ്പാടും അറിയിക്കുന്നു.  
പ്രേമ. എസ്  
മനു ജോസഫ്  
ഭാഷാശാസ്ത്രവിഭാഗം കേരളസർവകലാശാല.

## 2 ആവൃത്തി കണ്ടെത്തൽ

### 2.1 നിർവചനവും വിശദീകരണവും:

വർഗ്ഗീകരിക്കപ്പെട്ട ഡേറ്റയുടെ ക്രമീകൃതമായ വിന്യാസത്തെയാണ് ആവൃത്തി വിതരണം എന്നു പറയുന്നത്. ഡേറ്റയെ ഉചിതമായ രീതിയിൽ തരംതിരിച്ച് അവയുടെ വിതരണത്തിന്റെ എണ്ണം സൂചിപ്പിക്കുമ്പോൾ ആവൃത്തി വിതരണം ലഭ്യമാകും. യഥാർത്ഥത്തിൽ ഡാറ്റയെ സംഗ്രഹിയ്ക്കലാണ് ആവൃത്തിവിതരണത്തിലൂടെ ചെയ്യുന്നത്.

ആവൃത്തി കണ്ടെത്തൽ എന്നതുകൊണ്ട് ഉദ്ദേശിക്കുന്നത് ഒരു ഭാഷയിൽ ഓരോ ലിപിയും എത്ര തവണ ഉപയോഗിക്കപ്പെടുന്നു എന്ന് കണക്കാക്കുന്നതാണ്. അച്ചടിയിലും എഴുത്തിലും ഇവ എത്ര ശതമാനം ഉപയോഗിക്കപ്പെടുന്നു എന്ന ആവൃത്തി വിതരണകണക്കിന്റെ അടിസ്ഥാനത്തിലാണ് ഏറ്റവും കൂടുതൽ വരുന്ന അക്ഷരങ്ങളും, അവ ഏറ്റവും കൂടുതൽ ആവൃത്തിയിൽ ഉപയോഗിക്കുന്നത് ഏതെന്ന് സ്ഥാനങ്ങളിലാണെന്നും കണ്ടെത്തുന്നത്.

### 2.2 മുൻകാല പഠനങ്ങൾ

മലയാളം ടൈപ്പറൈറ്റിന്റെ കീ ബോർഡ് സംവിധാനം പരിഷ്കരിക്കുന്നതിന് വേണ്ടി 1971 - ൽ കേരള ഭാഷാ ഇൻസ്റ്റിറ്റ്യൂട്ടിന്റെ ആഭിമുഖ്യത്തിൽ ശ്രീ. എം. ആർ. വിജയനാഥൻപിള്ള നടത്തിയതാണ് ഈ മേഖലയിലെ ആദ്യപഠനം. ഇതിന്റെ അടിസ്ഥാനത്തിലാണ്, അഞ്ഞൂറിൽപ്പരം ലിപികൾ ഉപയോഗിച്ച് അച്ചടിച്ചുവന്നിരുന്ന മലയാളഭാഷയെ അക്കങ്ങളും ചിഹ്നങ്ങളും ഉൾപ്പെടെ പരമാവധി 92 അച്ചുകളിൽ ഒതുക്കിയത്.

92 അച്ചുകളിൽ 25 എണ്ണം അക്കങ്ങൾക്കും ചിഹ്നങ്ങൾക്കും മാറ്റിവെച്ച് ബാക്കി 67 സ്ഥാനങ്ങളിൽ ലിപികൾ ഒതുക്കേണ്ടതായി വന്നു. ലിപി പരിഷ്കരണപദ്ധതി പ്രകാരം 90 ലിപികളിൽ 23 എണ്ണം അച്ചടിയിൽ നിലനിർത്തി. ഇത്തരം സന്ദർഭങ്ങളിലാണ് ആവൃത്തി വിതരണം കണ്ടെത്തുന്നതിന്റെ ആവശ്യകതയുണ്ടാകുന്നത്. സാങ്കേതികവും ശാസ്ത്രീയവുമായി ഏത് നിലനിർത്തണം, ഏത് ഉപേക്ഷിക്കണം എന്ന് നിശ്ചയിക്കുന്നത് ആവൃത്തി കണ്ടെത്തുന്നതിലൂടെ മാത്രമേ സാധ്യമാകൂ. ഉള്ള സ്ഥലത്ത് ഏറ്റവും അത്യാവശ്യമായവയെ വളരെ ഫലപ്രദമായ രീതിയിൽ സ്ഥാപിക്കണം. കീ ബോർഡ് സംവിധാനം പരിഷ്കരിക്കുന്നതിനുവേണ്ടി നടന്ന ഈ പഠനത്തിൽ അദ്ദേഹം 11781 ലിപികൾ അപഗ്രഥിക്കുകയും, അതിൽ ഏറ്റവും കൂടുതൽ ആവൃത്തിയിലുള്ള 66 ലിപികൾ സ്വീകരിക്കുകയും ചെയ്തു. ഉപേക്ഷിച്ച കൂട്ടക്ഷരങ്ങൾ ചന്ദ്രക്കലയിട്ട് പിരിച്ചെഴുതണം എന്ന തീരുമാനത്തിലുമെത്തി.

അദ്ദേഹത്തിന്റെ അപഗ്രഥനത്തിലൂടെ കണ്ടെത്തിയ ലിപികളുടെ ആവൃത്തിപ്പട്ടിക താഴെ കൊടുത്തിരിക്കുന്നു.

ലിപി	ശതമാനം
ു	8.02
ി	7.74
ം	7.66
ാ	7.21
ത	5.77
െ	3.51
ക	3.26
യ	3.13
ഓ	2.94
മ	2.89
വ	2.86
ര	2.83

ന	2.56
ഓ	2.36
പ	2.25
ല	2.25
ട	2.18
ച	2.00
ണ	1.83
ന്ന	1.81
സ	1.75
ക്ക	1.70
റ	1.68
ള	1.56
ബ	1.23
അ	1.22
ൻ	1.19
ൽ	0.90
ർ	0.89
ു	0.82
ദ	0.81
ത	0.67
ദ	0.67
ു	0.66
ീ	0.65
ശ	0.63
ു	0.58
ൾ	0.56
ച്ച	0.56
ഷ	0.51
ഇ	0.49
എ	0.49
ഗ	0.48
ധ	0.44
ഭ	0.43
ജ	0.41
ഖ	0.37
ഴ	0.36
ഹ	0.33
ഥ	0.28
ബ	0.25
ഉ	0.20
ഈ	0.20
ൂ	0.17
ൗ	0.15
ഡ	0.13
ഖ	0.12
ൺ	0.12

ഏ	0.09
ഫ	0.08
ഘ	0.05
ഓ	0.05
ഛ	0.01
ഝ	0.003
ഞ	0.002
ത	0.000

(മാതൃഭൂമി ആഴ്ചപ്പതിപ്പ് 1971 ഒക്ടോബർ 3 പുസ്തകം 49 ലക്കം 29. പുതിയ മലയാളം അച്ചെഴുത്തുപദ്ധതി ഡോ.കെ. ഭാസ്കരൻ നായർ)

ഈ കണ്ടെത്തലിന്റെ അടിസ്ഥാനത്തിൽ വിരലുകളുടെ പ്രയോഗ ലഘുവത്തെയും അക്ഷരങ്ങളുടെ സ്ഥാനങ്ങൾ ഓർമ്മിക്കാനുള്ള സൗകര്യത്തെയും മുൻനിർത്തി ഭാഷാ ഇൻസ്റ്റിറ്റ്യൂട്ടിലെ ഡോക്ടർ എം. പി. പരമേശ്വരൻ ആദ്യം സംവിധാനം ചെയ്യുകയും പിന്നീടിതിൽ അച്ചെഴുത്തു കമ്മറ്റി ചില ഭേദഗതികൾ വരുത്തി കീബോർഡ് സംവിധാനം പരിഷ്കരിക്കാൻ സ്വീകരിക്കുകയും ചെയ്തു..

1994-ൽ സെൻട്രൽ ഇൻസ്റ്റിറ്റ്യൂട്ട് ഓഫ് ഇന്ത്യൻ ലാംഗ്വേജസ്, മൈസൂർ, മലയാളത്തിലെ സ്വനിമങ്ങളുടേയും രൂപിമങ്ങളുടേയും ആവൃത്തി കണ്ടെത്തി.. പുന ഡക്കാൻ കോളേജിലെ ഡോ.എം.എൻ ആപ്തെയുടേയും ഡോ. എം. എം. ഘാടേജിയുടെയും മേൽനോട്ടത്തിൽ ശ്രീ. കെ.എം.എൻ മേനോൻ, നടത്തിയ പഠനമാണ് 'ഫോണിമിക് ആന്റ് മോർഫിമിക് ഫ്രീക്വൻസി കൗണ്ട് ഇൻ മലയാളം' (Phonemic and Morphemic Frequency Count in Malayalam) എന്ന പുസ്തകം. മലയാളത്തിലെ എഴുത്തിലും സംസാരത്തിലുമുള്ള സ്വനിമങ്ങൾ, പദാംശങ്ങൾ, രൂപിമങ്ങൾ, വാക്കുകൾ തുടങ്ങിയവയുടെ പുനരാഗമനത്തിന്റെ ആവൃത്തി കണ്ടെത്തി പ്രത്യേക രീതിയിൽ ക്രമീകരിച്ചു. മലയാളത്തിൽ 'ഷോർട്ട് ഹാൻഡ്' സംവിധാനം കൊണ്ടുവരുന്നതിനു വേണ്ടിയാണ് ഈ പഠനം നടത്തിയത്. കൂടാതെ വിദേശ വിദ്യാർത്ഥികളുടെ ഭാഷാപഠനത്തെ സഹായിക്കുന്ന തരത്തിൽ ഏറ്റവും കൂടുതലായുള്ള വാക്കുകൾ ക്രമീകരിച്ച് പട്ടിക തയ്യാറാക്കി. മലയാള സാഹിത്യത്തിന്റെ വിവിധ മേഖലകളിൽ നിന്ന് (അതായത്, നോവലുകൾ, ചെറുകഥകൾ, കവിതകൾ, ലേഖനങ്ങൾ, ആനുകാലികങ്ങൾ, റേഡിയോ നാടകങ്ങൾ, സ്കൂൾ പാഠപുസ്തകങ്ങൾ, ശാസ്ത്രപുസ്തകങ്ങൾ, സാഹിത്യവിമർശനങ്ങൾ തുടങ്ങിയവ) ഡാറ്റാ സമാഹരണം നടത്തി. 100,000 വാക്കുകൾ സാംപ്ലിംഗ് രീതിയിലൂടെ ശേഖരിച്ച് അപഗ്രഥനം നടത്തി. ഏകദേശം മൂന്ന് വർഷത്തോളം നീണ്ടുനിന്ന ഈ പ്രവർത്തനത്തിൽ മുഴുവൻ വാക്കുകളും അകാരാദിക്രമത്തിൽ അവയുടെ ആവൃത്തിയോടൊപ്പം പട്ടികയിലാക്കി. സ്വനിമങ്ങളിൽ ആവൃത്തി ഏറ്റവും കൂടുതൽ വരുന്ന സ്വരാക്ഷരങ്ങൾ 'അ' 'ഉ' 'ഊ' എന്നിവയും സർവ്വസാധാരണമായി ആഗമിക്കുന്ന വ്യഞ്ജനാക്ഷരങ്ങൾ 'മ' 'ന' 'ക' തുടങ്ങിയവയുമാണെന്ന് കണ്ടെത്തുകയുണ്ടായി. രൂപിമങ്ങളിൽ ഏറ്റവും കൂടുതൽ ആഗമിച്ചിരിക്കുന്നത് പരപ്രത്യയം (Postposition) ആയ 'ഉം' ആണ്. രണ്ടാമതായി, വിഭക്തി രൂപിമങ്ങളും, അടുത്തത് കാലസൂചിക പരപ്രത്യയങ്ങളുമാണ്.

**2.3 ഈ പഠനത്തിന്റെ ഉദ്ദേശ്യങ്ങൾ - ഉപയോഗങ്ങൾ**

മുൻകാല പഠനങ്ങളെ അപേക്ഷിച്ച് ഈ പഠനം വളരെയധികം പ്രാധാന്യം അർഹിക്കുന്നു. അവ രണ്ടും രണ്ട് വ്യത്യസ്തങ്ങളായ ഉദ്ദേശ്യത്തോടെ നടന്നിട്ടുള്ളതാണെങ്കിലും വിവിധോദ്ദേശ്യത്തോടുകൂടിയ ഒരു പഠനം ഈ കാലഘട്ടത്തിൽ നടക്കേണ്ടതുണ്ട്. പ്രത്യേകിച്ചും മലയാളഭാഷയിലുള്ള വസ്തുതകളുടെ ശേഖരണത്തിനും വിശകലനത്തിനും കമ്പ്യൂട്ടർ സങ്കേതങ്ങൾ വ്യാപകമായി ഉപയോഗിച്ചുവരുന്ന സാഹചര്യത്തിൽ.

ഈ പഠനത്തിന്റെ ഉദ്ദേശ്യങ്ങൾ താഴെ പറയുന്നവയാണ്.

1. മലയാളം അക്ഷരങ്ങളുടെ (letters) ആവൃത്തി കണ്ടെത്തുക.
2. മലയാളം ക്യാരക്ടറുകളുടെ (characters) ആവൃത്തി കണ്ടെത്തുക.

3. ഓരോ അക്ഷരവും ആദ്യം, മധ്യം, അന്ത്യം എന്നിവിടങ്ങളിൽ വരുന്നതിന്റെ ആവൃത്തി കണ്ടെത്തുക.

4. ഏറ്റവും കൂടുതൽ ആവർത്തിക്കുന്ന വാക്കുകൾ കണ്ടെത്തുക.

ഈ ഗവേഷണകണ്ടെത്തലുകളുടെ ഉപയോഗങ്ങൾ താഴെ പറയുന്നു.

1. ഭാഷാ സാങ്കേതികവിദ്യയിലെ ഗവേഷണത്തിന്റെ ആദ്യപടിയെന്ന നിലയിൽ നടത്തിയ ഈ പഠനം വരും കാലങ്ങളിൽ കൂടുതൽ പഠന ഗവേഷണങ്ങൾക്ക് ഒരു മുതൽക്കൂട്ടായിരിക്കും.

2. മലയാള അക്ഷരങ്ങളുടെ ആവൃത്തി കണ്ടെത്തിയതിലൂടെ ഏറ്റവും ഫലപ്രദമായതും സൗകര്യപ്രദവുമായ രീതിയിൽ മലയാളം കീ ബോർഡ് സംവിധാനം പുനഃക്രമീകരിക്കാം.

ഉദാഹരണമായി ഏറ്റവും കൂടുതൽ ആവൃത്തി ഉപയോഗിക്കപ്പെടുന്ന ചില്ലുകൾക്കും കൂട്ടക്ഷരങ്ങൾക്കും രണ്ടിലധികം കീ ഉപയോഗിക്കേണ്ടിവരുന്നത് ടൈപ്പ് റൈറ്റിംഗ് വിരസവും വിഷമകരവുമാക്കിത്തീർക്കുന്നു.

3. മലയാള ഭാഷാപ്രചാരണത്തിന്റെ ഭാഗമായി വിദേശികൾക്കും ഒപ്പം വിദേശങ്ങളിൽ വസിക്കുന്ന മലയാളികളുടെ കുട്ടികൾക്കും ഭാഷാപഠനം എളുപ്പമാക്കിത്തീർക്കുന്നതിന്. ഏറ്റവും കൂടുതൽ ആവർത്തിക്കുന്ന പദങ്ങൾ പരിചയപ്പെടുത്തുന്നതിലൂടെ അടിസ്ഥാന പദസമ്പത്ത് വർദ്ധിപ്പിക്കാൻ സാധിക്കും.

4. ഇത്തരത്തിലുള്ള പഠനങ്ങൾ മറ്റ് ഭാഷകളിൽ കൂടി നടത്തി പദങ്ങളുടെ താരതമ്യപഠനത്തിന് സാധ്യതയുണ്ടാക്കാം.

### 2.4 ഡാറ്റ

1. മാതൃഭൂമി, മലയാളമനോരമ എന്നീ പത്രങ്ങളുടെ ഒരു മാസത്തെ പേജുകൾ.

2. മലയാളത്തിൽ ഏറ്റവും കൂടുതൽ അച്ചടിയ്ക്കപ്പെട്ട 'രമണൻ' 'ഖസാക്കിന്റെ ഇതിഹാസം' എന്നീ കൃതികൾ.

3. കേരള സർവ്വകലാശാല മലയാള വിഭാഗത്തിലെ എം.എ. മലയാളം പരീക്ഷയുടെ നൂറ് ഉത്തരക്കടലാസുകൾ.

#### 2.4.1 ഡേറ്റ ശേഖരിക്കുന്നതിന് തെരഞ്ഞെടുത്ത രീതി.

1. അക്കങ്ങളിൽ രേഖപ്പെടുത്തിയിരിക്കുന്ന സംഖ്യകൾ, മലയാളം അക്ഷരത്തിലല്ലാതെ അച്ചടിച്ചിരിക്കുന്ന വാക്കുകൾ തുടങ്ങിയവ ഒഴിവാക്കിയിട്ടുണ്ട്.

2. ഡാറ്റാ ശേഖരിച്ചത് വ്യവസ്ഥാപിതമായ സാംപ്ലിംഗ് രീതിയിലാണ്. മാതൃഭൂമി, മലയാളമനോരമ എന്നീ പത്രങ്ങളുടെ ഒരു മാസത്തെ പത്രത്തിൽ നിന്ന് എല്ലാ ശനിയാഴ്ചകളിലെയും ഞായറാഴ്ചകളിലെയും പേജുകൾ.

3. 'ഖസാക്കിന്റെ ഇതിഹാസം' 'രമണൻ' എന്നീ കൃതികളിൽ നിന്ന് റാൻഡം സാംപ്ലിംഗിലൂടെ തെരഞ്ഞെടുത്ത പേജുകൾ.

4. എം.എ മലയാളം പരീക്ഷയുടെ നൂറ് ഉത്തരക്കടലാസുകളിൽ നിന്ന് സാംപ്ലിംഗ് രീതിയിൽ തെരഞ്ഞെടുത്ത പേജുകൾ.

### 2.5 ഗവേഷണരീതി

മലയാളത്തിലുള്ള വസ്തുതകൾ കമ്പ്യൂട്ടറിലേക്ക് പ്രവേശിപ്പിക്കുവാൻ ഇന്ന് വ്യാപകമായി ഉപയോഗിക്കുന്ന ഐ.എസ്.എം മലയാളം എന്ന സോഫ്റ്റ്‌വെയറിൽ തയ്യാറാക്കിയ വിവരങ്ങൾ പ്രത്യേകമായി വികസിപ്പിച്ച പ്രോഗ്രാമുപയോഗിച്ച് വിശകലനം ചെയ്താണ് അക്ഷരങ്ങളുടെ ആവൃത്തി, വിതരണക്രമം, വാക്കുകളുടെ ആവൃത്തി, എന്നിവ കണ്ടുപിടിച്ചത്.

പ്രത്യേകമായ ഡിജിറ്റൽ കോഡുകൾ ഉപയോഗിച്ചാണ് കമ്പ്യൂട്ടർ അക്ഷരങ്ങളെയും ചിഹ്നങ്ങളെയും ആന്തരികമായി (അതായത് മെമ്മറിയിൽ) പ്രതിനിധീകരിക്കുന്നത്. ആസ്കി, ഇസ്കി, യൂണികോഡ് എന്നിവയെല്ലാം ഇങ്ങനെ അക്ഷരങ്ങൾക്കും ചിഹ്നങ്ങൾക്കുമായി ഡിജിറ്റൽ കോഡുകൾ ഉണ്ടാക്കാൻ ഉപയോഗിക്കുന്ന സമ്പ്രദായങ്ങളാണ്. എന്നാൽ ഐ.എസ്.എം സോഫ്റ്റ്‌വെയറിന്റെ മലയാളം പതിപ്പ് വ്യത്യസ്തമായ ഒരു രീതിയാണ് വിവരപ്രതിനിധീകരണത്തിനായി ഉപയോഗിക്കുന്നത്.

സ്ക്രീനിൽ കാണിക്കുന്ന ഓരോ ചിഹ്നത്തിനും തുല്യമായി കോഡുകൾ നൽകിയാണ് അക്ഷരങ്ങളെ കമ്പ്യൂട്ടറിൽ പ്രതിനിധീകരിക്കുന്നത്. അതിനാൽ തന്നെ അക്ഷരകോഡുകളെ അടിസ്ഥാനമാക്കി മലയാളത്തിലുള്ള വിവരങ്ങളെ വിശകലനം ചെയ്യുക സാധ്യമല്ല. ഉദാഹരണമായി 'ഇ' എന്ന അക്ഷരത്തിന് തുല്യമായ കോഡ് 67 ആണ്. അതേ സമയം 'ഈ' എന്ന അക്ഷരം പ്രതിനിധീകരിക്കപ്പെടുന്നത് 'ഇ' എന്ന അക്ഷരവും 'ഊ' എന്ന ചിഹ്നവും കൂടിച്ചേർന്ന രൂപമായാണ്. മറ്റൊരു വിധത്തിൽ പറഞ്ഞാൽ, 'ഇ' യുടെയും 'ഊ' യുടെയും കോഡുകൾ കൂടിച്ചേർന്നാണ് 'ഈ' കമ്പ്യൂട്ടറിൽ പ്രതിനിധീകരിക്കപ്പെടുന്നത്. ഇസ്ഫോക് (ISFOC) എന്ന ഫോണ്ട് കോഡിംഗ് സമ്പ്രദായം ഉപയോഗിച്ചാണ് ഐ.എസ്.എം പ്രവർത്തിക്കുന്നത്.

ഇസ്ഫോക് കോഡിംഗ് സമ്പ്രദായത്തെ അടിസ്ഥാനമാക്കി പ്രവർത്തിക്കുന്ന പ്രോഗ്രാം തയ്യാറാക്കി ടെക്സ്റ്റ് ഫയലുകളുടെ രൂപത്തിലുള്ള വിവരങ്ങളിൽ നിന്ന് മലയാള അക്ഷരങ്ങളെ വേർതിരിച്ചെടുത്ത് എണ്ണിയാണ് ആവൃത്തി കണ്ടെത്തിയത്.

രണ്ട് വീക്ഷണകോണിൽ നിന്നുകൊണ്ട് മലയാളത്തിലെ അക്ഷരങ്ങളുടെ തരംതിരിവ് നടത്തിയാണ് ആവൃത്തി നിർണ്ണയം നടത്തിയത്.

1. സ്വരങ്ങൾ, വ്യഞ്ജനങ്ങൾ, കൂട്ടക്ഷരങ്ങൾ, ചില്ലുകൾ, ചിഹ്നങ്ങൾ എന്നിവയുടെ ഒറ്റക്കൊറ്റക്കുള്ള ആവൃത്തി (ഉദാഹരണമായി 'ആവൃത്തി' എന്ന വാക്ക് 'ആ', 'വ', 'ൂ', 'ത്ത,' 'ി' എന്നീ അക്ഷരങ്ങൾ ചേർന്നതായി കണക്കാക്കിക്കൊണ്ട്)
2. ചിഹ്നങ്ങളും അക്ഷരങ്ങളും കൂടിച്ചേർന്ന രൂപങ്ങളെ അക്ഷരമായി കണക്കാക്കിക്കൊണ്ടുള്ള നിർണ്ണയം. വാക്കുകളിലെ അക്ഷരങ്ങളുടെ വിന്യാസം കണ്ടുപിടിച്ചത് ഈ രീതിയാലാണ്.

അക്ഷരങ്ങളുടെ ഇടയിൽ വരുന്ന ശൂന്യസ്ഥലങ്ങൾ, ടാബുകൾ, കുത്തുകൾ, ആശ്ചര്യ ചിഹ്നങ്ങൾ എന്നിവ വാക്കുകളുടെ അതിർത്തി നിർണ്ണയിക്കാൻ ഉപയോഗിച്ചു.

### 3 ആവൃത്തിപ്പട്ടിക

#### 3.1 അക്ഷരാടിസ്ഥാനത്തിൽ(letters)

മലയാളഅക്ഷരങ്ങളിൽ ഏറ്റവും കൂടുതൽ ആവൃത്തിയിൽ വരുന്നത് 'ി' (9.19%) എന്ന സ്വര ചിഹ്നവും 'ക' (3.49%) എന്ന വ്യഞ്ജനാക്ഷരവുമാണ്.

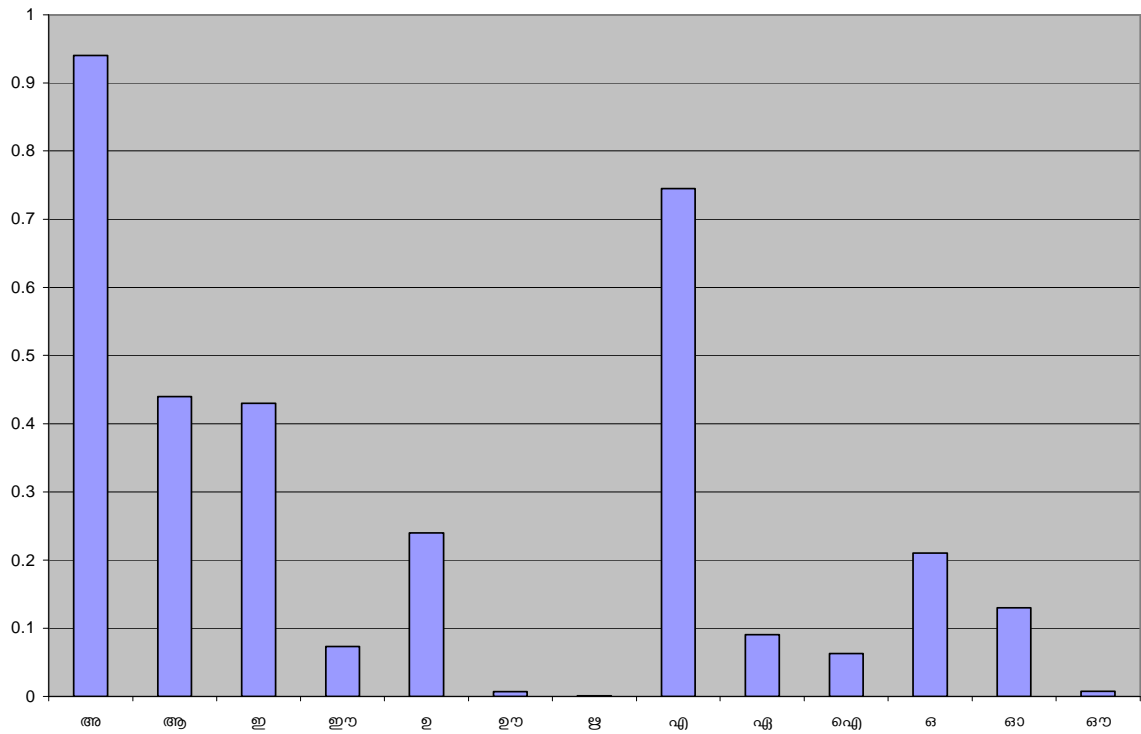
##### 3.1.1 സ്വരാക്ഷരങ്ങൾ (Vowels)

സ്വരാക്ഷരങ്ങളിൽ ഏറ്റവും കൂടുതൽ ആവൃത്തിയുള്ളത് അ -0.94%.  
 ഏറ്റവും കുറവ് ഈ -0.0067%..

ആകെ പരിശോധിച്ച അക്ഷരങ്ങൾ 1520868



അക്ഷരം	ആവൃത്തി	ശതമാനം
അ	14311	0.94
എ	11366	0.745
ആ	6724	0.44
ഇ	6539	0.43
ഉ	3691	0.24
ഒ	3258	0.21
ഓ	1933	0.13
ഏ	1382	0.091
ഈ	1109	0.073
ഐ	959	0.063
ഔ	115	0.0076
ഊ	102	0.0067
ഘ	13	0.00085

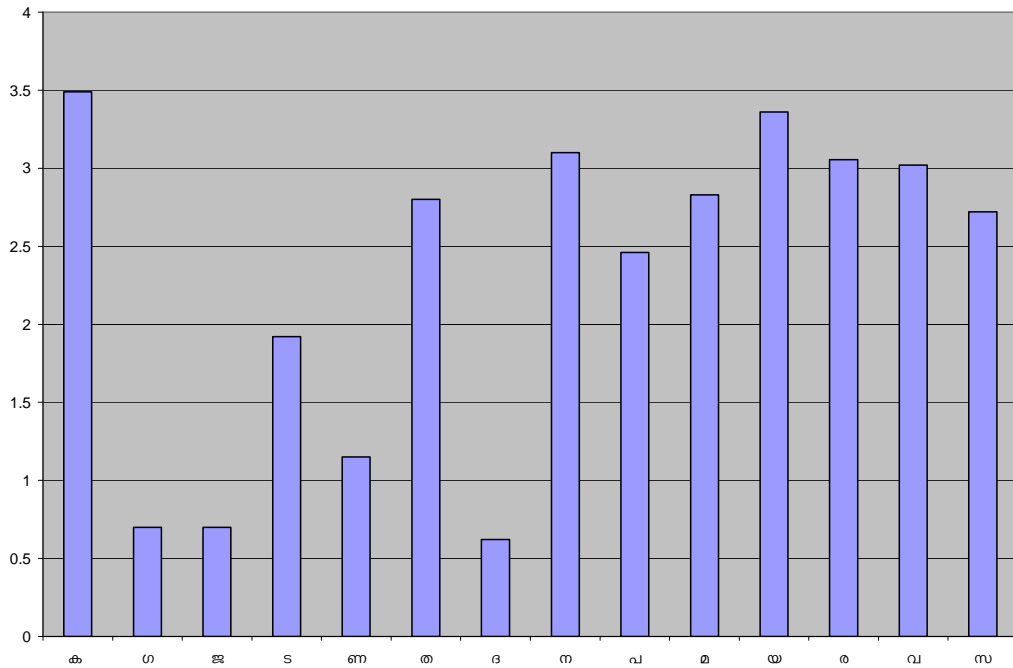


### 3.1.2 വ്യജ്ഞനാക്ഷരങ്ങൾ (Consonants)

വ്യജ്ഞനാക്ഷരങ്ങളിൽ ഏറ്റവും കൂടുതൽ ആവൃത്തിയുള്ളത് ക -3.49%.  
 ഏറ്റവും കുറവ് തയ -0.00078%.

ആകെ പരിശോധിച്ച അക്ഷരങ്ങൾ 1520868

അക്ഷരം	ആവൃത്തി	ശതമാനം
ക	53088	3.49
യ	51155	3.36
ന	47153	3.10
ര	46469	3.055
വ	45964	3.02
മ	42978	2.83
ത	42772	2.8
സ	41359	2.72
പ	37563	2.46
ട	29255	1.92
ല	26390	1.74
റ	19509	1.28
ണ	17570	1.15
ള	17082	1.12
ഗ	10640	0.70
ശ	10617	0.70
ജ	10606	0.70
ദ	9439	0.62
ഷ	9151	0.60
ബ	8925	0.59
ച	8780	0.58
ഡ	8019	0.53
ഹ	7194	0.47
ഭ	6811	0.45
ഴ	6582	0.43
ഫ	5127	0.34
ധ	5058	0.33
ഖ	2192	0.14
ഘ	1533	0.10
ഘ	987	0.065
ഞ	723	0.048
ഠ	698	0.046
ണ	649	0.043
ഡ	55	0.0036
ഛ	38	0.0025
തയ	12	0.00078

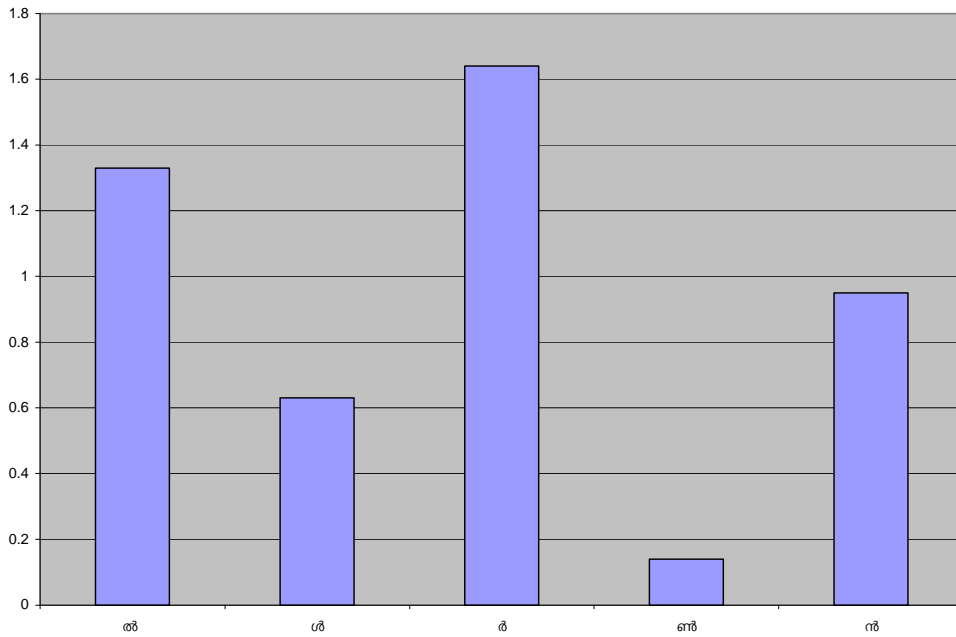


### 3.1.3 ചില്ലുകൾ

ചില്ലക്ഷരങ്ങളിൽ ഏറ്റവും കൂടുതൽ ആവൃത്തിയുള്ളത് ൽ -1.64%.  
 ഏറ്റവും കുറവ് ണ് -0.14%.

ആകെ പരിശോധിച്ച അക്ഷരങ്ങൾ 1520868

അക്ഷരം	ആവൃത്തി	ശതമാനം
ൽ	20226	1.33
ൾ	9647	0.63
ർ	24929	1.64
ൺ	2195	0.14
ൻ	14390	0.95



### 3.1.4 കൂട്ടക്ഷരങ്ങൾ (Consonant clusters)

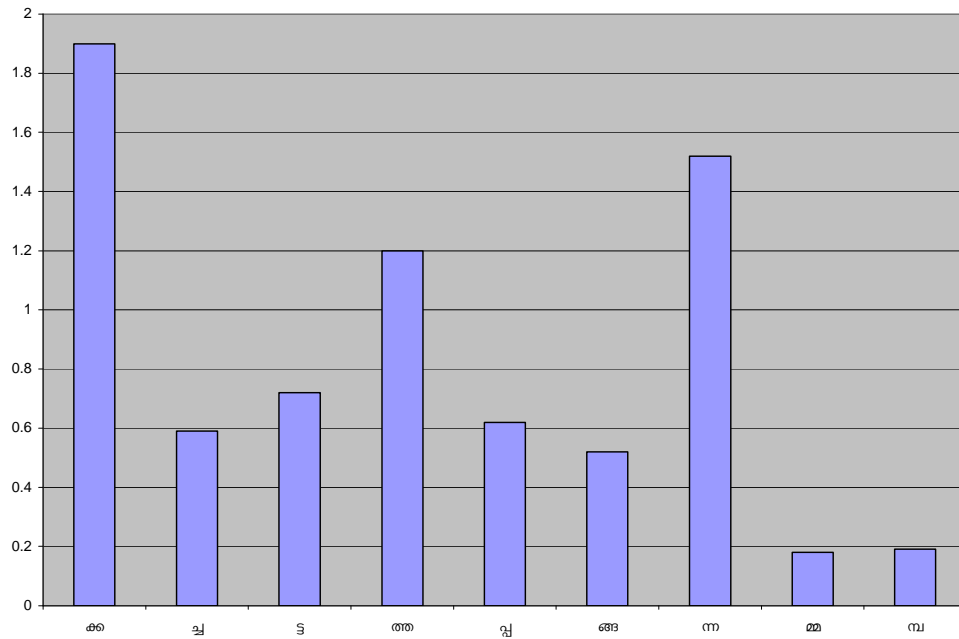
കൂട്ടക്ഷരങ്ങളിൽ ഏറ്റവും കൂടുതൽ ആവൃത്തിയുള്ളത് ക്ക -1.90%.

ഏറ്റവും കുറവ് ണ്ട്ട -0.00059%.

ആകെ പരിശോധിച്ച അക്ഷരങ്ങൾ 1520868

അക്ഷരം	ആവൃത്തി	ശതമാനം
ക്ക	28881	1.90
ന്ന	23059	1.52
ത്ത	18699	1.2
ട്ട	10970	0.72
പ്പ	9502	0.62
ച്ച	9011	0.59
ങ്ങ	7922	0.52
ണ്ട	7379	0.49
ന്റ	6074	0.40
റ്റ	5175	0.34
ന്ത	5065	0.33
ല്ല	5050	0.33
ക്ഷ	3899	0.26
ഞ്ഞ	3082	0.20
ള്ള	2934	0.19
മ്പ	2911	0.19
മ്മ	2777	0.18
ങ്ക	2527	0.17
സ്ഥ	2103	0.14
ന്ദ	2050	0.13
സ്സ	1901	0.12

ബ	1473	0.097
യ	1427	0.094
ല	1351	0.089
സ	1049	0.069
സ്	1047	0.069
ണ്	1023	0.067
ദ	879	0.058
ക	876	0.058
ത	477	0.03
സ	447	0.029
ശ	366	0.024
മ	364	0.024
മ്മ	255	0.017
വ	249	0.016
ജ	147	0.0097
ജ്	129	0.0085
ശ്	124	0.0082
ച്	96	0.0063
ബ്	88	0.0058
ശ്	83	0.005
ഫ്	78	0.0051
ന	74	0.0049
ഗ്	51	0.0033
ഭ്	37	0.0024
ഹ്	18	0.0012
ബ്	15	0.00098
ശ്	13	0.00085
ഡ്	9	0.00059

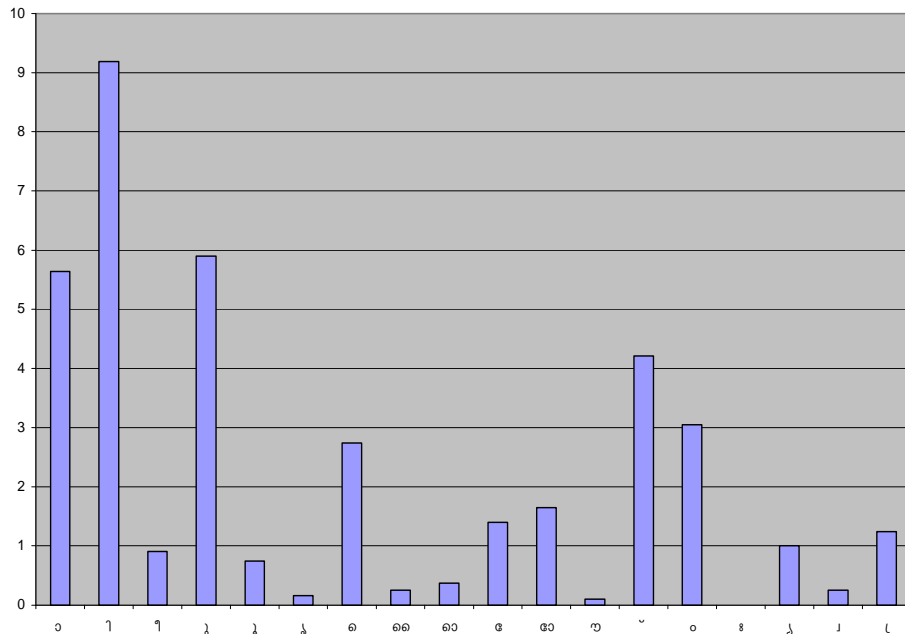


### 3.1.5 ചിഹ്നാടിസ്ഥാനത്തിൽ (Allographic)

ചിഹ്നങ്ങളിൽ ഏറ്റവും കൂടുതൽ ആവൃത്തിയുള്ളത് ി -9.19%.  
 ഏറ്റവും കുറവ് ഖ -0.0022%.

ആകെ പരിശോധിച്ച അക്ഷരങ്ങൾ 1520868

അക്ഷരം	ആവൃത്തി	ശതമാനം
ി	139814	9.19
ു	89800	5.90
ഊ	85839	5.64
ഊ	64064	4.21
ഊ	46418	3.05
ഊ	41620	2.74
ഊ	25062	1.65
ഊ	21948	1.4
ഊ	18909	1.24
ഊ	15228	1.00
ഊ	13794	0.91
ഊ	11300	0.74
ഊ	5609	0.37
ഊ	3787	0.25
ഊ	3773	0.25
ഊ	2423	0.16
ഊ	1591	0.10
ഊ	33	0.0022



### 3.2 താരതമ്യം

1971-ൽ നടത്തിയ പഠനവും നിലവിൽ നടത്തിയ പഠനവും താരതമ്യം ചെയ്യുമ്പോൾ മലയാള അക്ഷരങ്ങളുടെ ആവൃത്തിയിൽ പ്രകടമായ വ്യത്യാസങ്ങൾ കാണാം. ഭാഷ സ്ഥിരസ്വഭാവം പ്രകടിപ്പിക്കുന്ന പ്രതിഭാസമല്ല, മറിച്ച് കാലത്തിനൊത്ത് മാറിക്കൊണ്ടിരിക്കുന്നു എന്ന തത്വത്തിന് ഇത്തരമൊരു താരതമ്യം അടിവരയിടുന്നു. എങ്കിൽ തന്നെയും അടിസ്ഥാനപരമായി ചില അക്ഷരങ്ങൾ തന്നെയാണ് കൂടുതലായി മലയാളത്തിൽ ഈ കാലയളവിൽ ഉപയോഗിക്കപ്പെടുന്നതെന്ന് കാണാൻ കഴിയും. 1971-ലെയും 2004-ലെയും പഠനങ്ങളിൽ 'അ' തന്നെയാണ് സ്വരവിഭാഗത്തിൽ ഏറ്റവും കൂടുതൽ ആവൃത്തിയുള്ള അക്ഷരം. അതേസമയം 1971-ൽ 'ഏ' ആയിരുന്നു ഏറ്റവും കുറവെങ്കിൽ 2004-ൽ അത് 'ഉ' ആണ്.

വ്യജ്ഞനങ്ങളിൽ 1971-ൽ 'ത' യാണ് ഏറ്റവും കൂടുതൽ ആവൃത്തിയുള്ള വന്നിട്ടുള്ള അക്ഷരം. 2004-ൽ അത് 'ക' യാണ്. 1971-ലും 2004-ലും ഏറ്റവും കുറവ് വന്നിട്ടുള്ളത് 'ഡ' യാണ്. പൊതുവെ 2004-ലും 1971-ലും കൂടുതലായി ആവൃത്തിയുള്ളവരുന്ന വ്യജ്ഞനാക്ഷരങ്ങൾ 'ക, ത, മ, യ, ര, വ, ന, സ, പ, ല, ട' എന്നിവയാണ്. കൂട്ടക്ഷരങ്ങളെപ്പറ്റിയുള്ള വിപുലമായ പഠനം 1971-ൽ നടന്നിട്ടില്ല. എങ്കിലും 1971-ലും 2004-ലും 'ക്ക' തന്നെയാണ് ഏറ്റവും കൂടുതൽ ആവൃത്തിയുള്ളവന്നിട്ടുള്ള കൂട്ടക്ഷരം. 'ക്ക, ന്ന, ത്ത, ട്ട, പ്പ, ച്ച, ണ്ണ, ണ്ട, ന്ന' എന്നിവയാണ് കൂടുതൽ ആവൃത്തിയുള്ള കൂട്ടക്ഷരങ്ങൾ. ചില്ലുകളുടെ കാര്യത്തിലും പ്രകടമായ വ്യത്യാസം കാണാം. 1971-ൽ 'ൽ' ആയിരുന്നു ഏറ്റവും കൂടുതൽ ആവൃത്തിയുള്ള ചില്ലെങ്കിൽ 2004-ൽ അത് 'ർ' ആണ്. ചിഹ്നങ്ങളിൽ 'ു' വും 'ി' യും ആണ് 1971-ലും 2004-ലും യഥാക്രമം ആവൃത്തിയുള്ളവരുന്നവ. 'ു, ി, റ, റ്' എന്നിവയാണ് കൂടുതൽ ആവൃത്തിയുള്ള ചിഹ്നങ്ങളായി രണ്ട് പഠനങ്ങളിലും കണ്ടെത്തിയിട്ടുള്ളത്.

2004-ൽ നടത്തിയ ഗവേഷണത്തിൽ പഠനത്തിനായെടുത്ത ഡാറ്റയുടെ അളവ് 1971-ലേതിനെ അപേക്ഷിച്ച് വളരെ കൂടുതലാണ്. 1971-ലെ പഠനത്തിൽ, മലയാളത്തിലെ എല്ലാ അക്ഷരങ്ങളും പരിഗണിച്ചിരുന്നില്ല (പ്രത്യേകിച്ച് കൂട്ടക്ഷരങ്ങൾ) അതിനാൽ തന്നെ ഓരോ അക്ഷരത്തിന്റെയും ആവൃത്തിയുടെ ശതമാനം വളരെ കൂടുതലാണ് (പട്ടിക നോക്കുക). മറ്റൊരു വ്യത്യാസം കമ്പ്യൂട്ടർ പ്രോഗ്രാമിന്റെ സഹായത്തോടെയാണ് നിലവിൽ നടത്തിയ പഠനം ഡാറ്റ വിശകലനം ചെയ്തത് എന്നുള്ളതാണ്. കമ്പ്യൂട്ടറുകളുടെ പ്രധാനമേന്മയായ കൃത്യതയാണ് ഇത്തരമൊരു രീതി തിരഞ്ഞെടുക്കാൻ പ്രേരിപ്പിച്ചത്.

### 3.3 ആവൃത്തിപ്പട്ടിക- അക്ഷരങ്ങളുകളുടെ വിതരണത്തിന്റെ അടിസ്ഥാനത്തിൽ

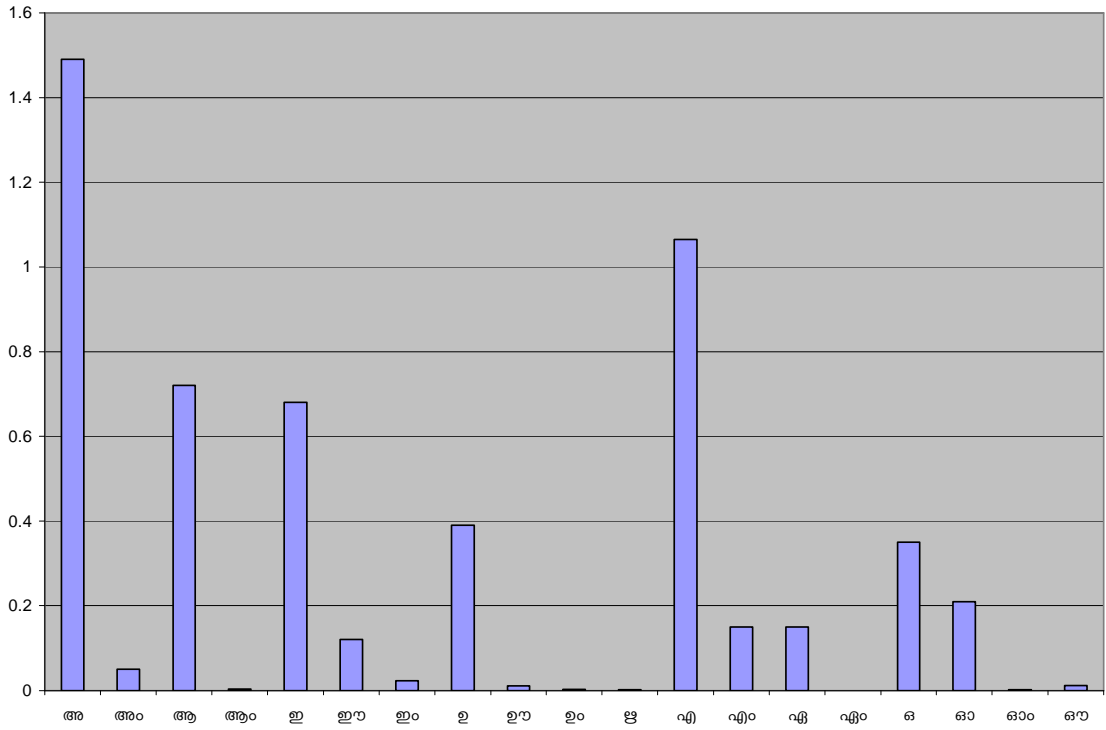
ചിഹ്നങ്ങളും ഒറ്റയ്ക്ക് നിലനിൽപുള്ള അക്ഷരങ്ങളും കുടിച്ചേർന്നരുപങ്ങളായി മലയാള അക്ഷരങ്ങളെ പരിഗണിച്ചാണ് (characters) അക്ഷരങ്ങളുടെ വാക്കുകളിലെ വിതരണം കണ്ടുപിടിച്ചത്. അക്ഷരങ്ങളുടെ ആവൃത്തി, വാക്കുകളുടെ തുടക്കം, മധ്യം, അവസാനം എന്നീ സ്ഥാനങ്ങളിൽ വരുന്നതിന്റെ ആവൃത്തി എന്നിവ യഥാക്രമം പട്ടികകളിൽ കൊടുത്തിരിക്കുന്നു. പരിശോധനയ്ക്ക് വിധേയമാക്കിയ വസ്തുതകൾ കമ്പ്യൂട്ടറിലേക്ക് പ്രവേശിപ്പിച്ചപ്പോഴുണ്ടായ തെറ്റുകൾ മൂലം ചില അക്ഷരങ്ങൾ അവ ഒരിക്കലും വരുവാൻ സാധ്യതയില്ലാത്ത സ്ഥാനങ്ങളിൽ അപൂർവ്വമായി വന്നിട്ടുണ്ട്.

#### 3.3.1 സ്വരാക്ഷരങ്ങൾ.

ആകെ പരിശോധിച്ച അക്ഷരങ്ങൾ 928971

അക്ഷരം	ആവൃത്തി	ആദൃം	മധ്യം	അന്ത്യം	ശതമാനം
അ	13811	13811	0	0	1.49
എ	9894	9894	0	0	1.065
ആ	6685	6685	0	0	0.72
ഇ	6321	6321	0	0	0.68
ഉ	3587	3587	0	0	0.39
ഒ	3255	3255	0	0	0.35
ഓ	1920	1920	0	0	0.21
ഏ	1379	1379	0	0	0.15
ഈ	1360	1360	0	0	0.15
ഊ	1109	1109	0	0	0.12
ഔ	464	464	0	0	0.050
ഘ	213	213	0	0	0.023
ഘ	115	115	0	0	0.012
ഘ	102	102	0	0	0.011
ഘ	28	28	0	0	0.0030
ഘ	21	21	0	0	0.002
ഘ	12	12	0	0	0.0013
ഘ	11	11	0	0	0.0012
ഘ	2	2	0	0	0.0002

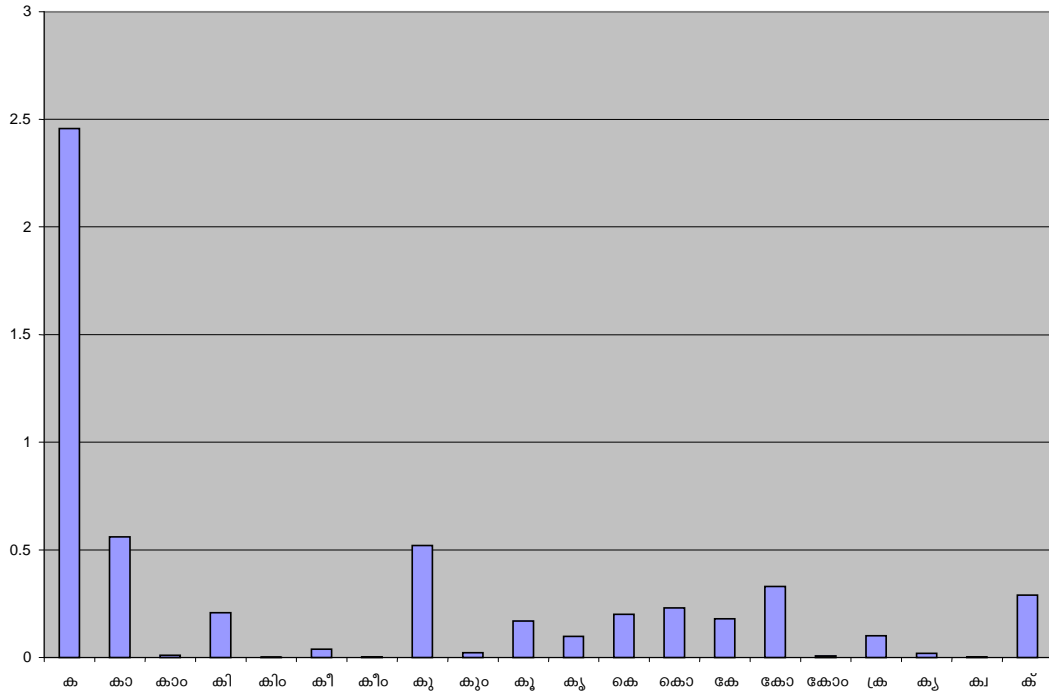




### 3.3.2 വ്യജ്ഞനാക്ഷരങ്ങൾ

ആകെ പരിശോധിച്ച അക്ഷരങ്ങൾ 928971

അക്ഷരം	ആവൃത്തി	ആദ്യം	മധ്യം	അന്ത്യം	ശതമാനം
ക	22813	6227	3958	2125	2.456
കാ	5178	2785	714	22	0.56
കാം	89	2	23	52	0.0096
കി	1923	760	416	227	0.207
കിം	12	10	0	1	0.0013
കീ	349	141	69	1	0.0376
കീം	26	0	0	26	0.0028
കു	4823	2252	936	15	0.52
കും	203	9	13	176	0.022
കൂ	1610	954	278	19	0.17
കൂ	916	349	291	0	0.099
കെ	1851	1664	29	108	0.20
കൊ	2109	1557	123	10	0.23
കേ	1674	1466	68	9	0.18
കോ	3100	2590	140	40	0.33
കോം	72	58	0	14	0.008
ക്ര	975	74	457	30	0.10
കൃ	175	50	51	12	0.019
കൃ	29	21	5	0	0.0031
ക്	2709	28	781	514	0.29



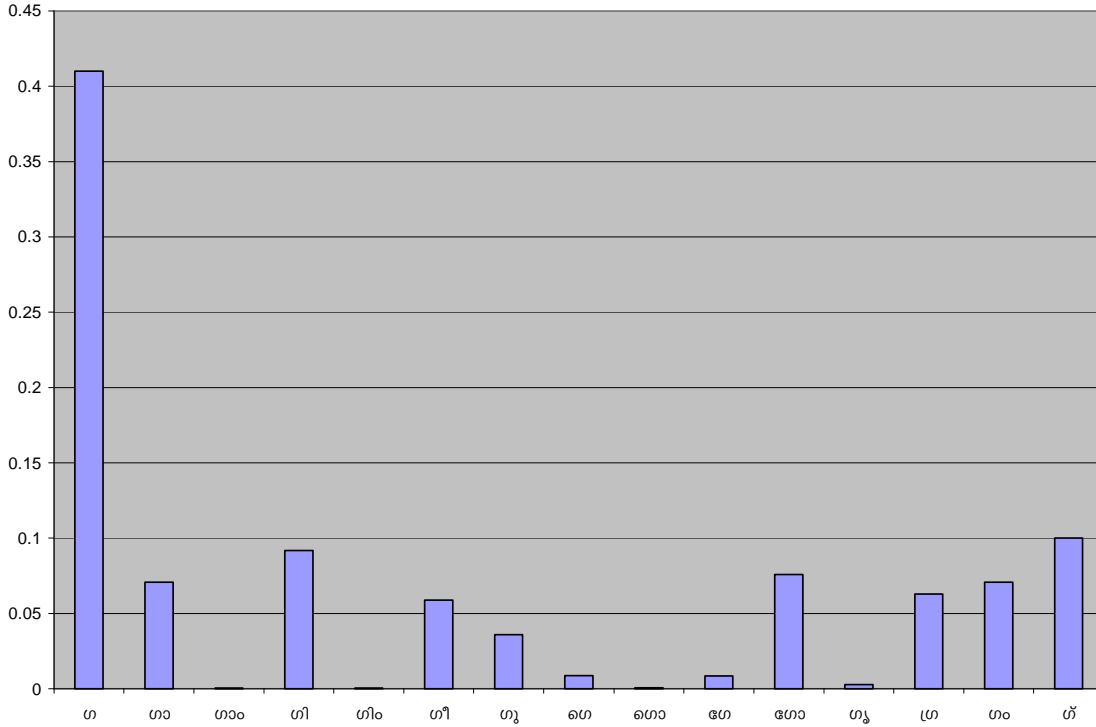
ആകെ പരിശോധിച്ച അക്ഷരങ്ങൾ 928971

അക്ഷരം	ആവൃത്തി	ആദ്യം	മധ്യം	അന്ത്യം	ശതമാനം
വ	1138	149	304	150	0.12250
വ്യ	347	0	129	43	0.0378
വി	157	3	73	8	0.017
വ്യാ	140	1	71	2	0.015
വാ	130	53	41	5	0.014
വേ	88	63	23	0	0.0095
വ്	63	0	3	55	0.0068
വം	61	0	0	61	0.0066
വു	16	6	8	1	0.0017
വോ	5	4	1	0	0.0005
വ്വ	3	0	0	3	.00033

ആകെ പരിശോധിച്ച അക്ഷരങ്ങൾ 928971

അക്ഷരം	ആവൃത്തി	ആദ്യം	മധ്യം	അന്ത്യം	ശതമാനം
ഗ	3773	566	1645	66	0.41
ഗ്	930	5	170	654	0.10
ഗി	852	73	267	31	0.092
ഗോ	703	490	116	32	0.076
ഗം	666	54	13	581	0.071
ഗാ	662	235	131	30	0.071
ഗ്ര	589	80	160	5	0.063

ഗീ	547	53	221	1	0.059
ഗു	330	241	43	0	0.036
ഗെ	83	75	6	2	0.0089
ഗേ	81	37	33	2	0.0087
ഗൂ	26	17	0	0	0.0028
ഗൊ	6	0	0	6	0.00065
ഗാം	5	4	0	1	0.0005
ഗിം	5	0	4	0	0.0005



ആകെ പരിശോധിച്ച അക്ഷരങ്ങൾ 928971

അക്ഷരം	ആവൃത്തി	ആദ്യം	മധ്യം	അന്ത്യം	ശതമാനം
ഘ	682	119	198	5	0.073
ഘാ	467	6	333	0	0.050
ഘാം	3	0	0	0	0.0003
ഘി	29	0	21	0	0.0031
ഘു	49	1	8	11	0.0053
ഘൂ	1	0	0	0	0.0001
ഘെ	2	2	0	0	0.0002
ഘോ	152	10	27	0	0.016
ഘ്	26	0	9	10	0.0028
ഘം	92	0	0	90	0.0010
ഘ്യ	10	0	10	0	0.0011
ഘ്യം	6	0	0	6	0.0006

ആകെ പരിശോധിച്ച അക്ഷരങ്ങൾ 928971

അക്ഷരം	ആവൃത്തി	ആദ്യം	മധ്യം	അന്ത്യം	ശതമാനം
ഒ	50	0	14	4	0.0054
ഒറ	2	0	2	0	0.0002
ഒറം	3	0	0	3	0.0003
ഒരി	16	0	5	1	0.0017
ഒരു	4	0	2	0	0.0004
ഒരും	4	0	0	4	0.0004
ഒർ	562	0	70	445	0.060
ഒരം	2	0	0	2	0.0002
ഒര്യ്യ	2	0	0	2	0.0002

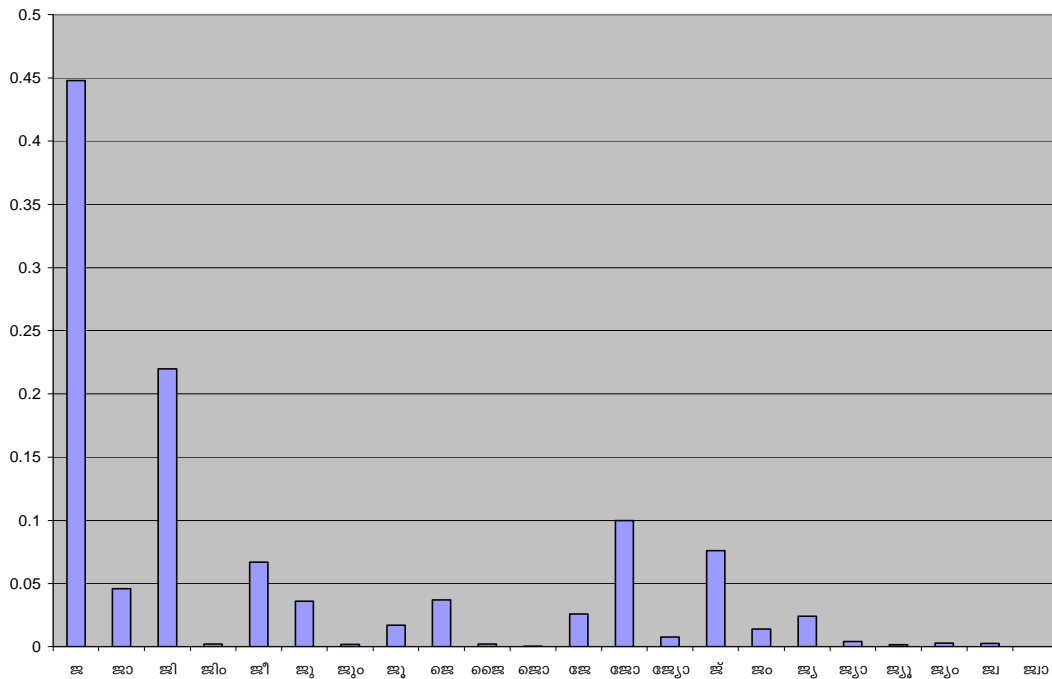
ആകെ പരിശോധിച്ച അക്ഷരങ്ങൾ 928971

അക്ഷരം	ആവൃത്തി	ആദ്യം	മധ്യം	അന്ത്യം	ശതമാനം
ച	2734	1017	507	413	0.29
ചറ	823	581	103	1	0.089
ചറം	55	53	0	0	0.0059
ചരി	1236	901	129	12	0.133
ചരീ	108	81	8	0	0.012
ചു	338	321	8	2	0.036
ചൂ	159	139	16	0	0.017
ചെ	2268	2147	54	2	0.244
ചൈ	43	43	0	0	0.004
ചെറ	78	74	0	0	0.008;
ചേ	444	343	28	0	0.0478
ചേറ	320	299	4	0	0.034
ചന്	14	14	0	0	0.0015
ചി	7	0	2	3	0.00075
ച്യ	24	0	9	3	0.0026

ആകെ പരിശോധിച്ച അക്ഷരങ്ങൾ 928971

അക്ഷരം	ആവൃത്തി	ആദ്യം	മധ്യം	അന്ത്യം	ശതമാനം
ഛ	8	5	2	0	0.0008
ഛറ	17	9	6	0	0.0018
ഛരി	5	1	3	0	0.0005
ഛർ	5	0	0	1	0.0005
ഛരം	2	0	0	0	0.0002
ഛെ	1	1	0	0	0.0001

അക്ഷരം	ആവൃത്തി	ആദൃം	മധ്യം	അന്ത്യം	ശതമാനം
ജ	4166	1755	988	187	0.448
ജാ	428	197	103	40	0.046
ജി	2058	925	317	299	0.22
ജിം	20	4	0	0	0.0022
ജീ	625	448	108	0	0.067
ജ്യ	332	32	139	88	0.036
ജ്യം	18	0	0	18	0.0019
ജ്ഞ	158	145	0	2	0.017
ജ്ഞ	342	320	14	4	0.037
ജ്ഞെ	20	17	2	0	0.0022
ജ്ഞൊ	2	2	0	0	0.00022
ജ്ഞേ	237	137	81	0	0.026
ജ്ഞോ	953	840	31	8	0.10
ജ്ഞ്യോ	70	59	0	0	0.0075
ജ്	705	0	145	505	0.076
ജം	127	92	7	25	0.014
ജ്യ	223	0	92	16	0.024
ജ്യൊ	39	0	37	0	0.0042
ജ്യൂ	15	1	0	0	0.0016
ജ്യം	27	0	2	25	0.0029
ജ്വ	25	23	2	0	0.0027
ജ്വാ	1	1	0	0	0.00010



ആകെ പരിശോധിച്ച അക്ഷരങ്ങൾ

928971

അക്ഷരം	ആവൃത്തി	ആദ്യം	മധ്യം	അന്ത്യം	ശതമാനം
ഡ	4	2	1	0	0.0004
ഡാ	2	0	0	0	0.0002
ഡി	4	0	0	0	0.0004
ഡ്യ	1	0	1	0	0.0001
ഡോ	1	1	0	0	0.0001

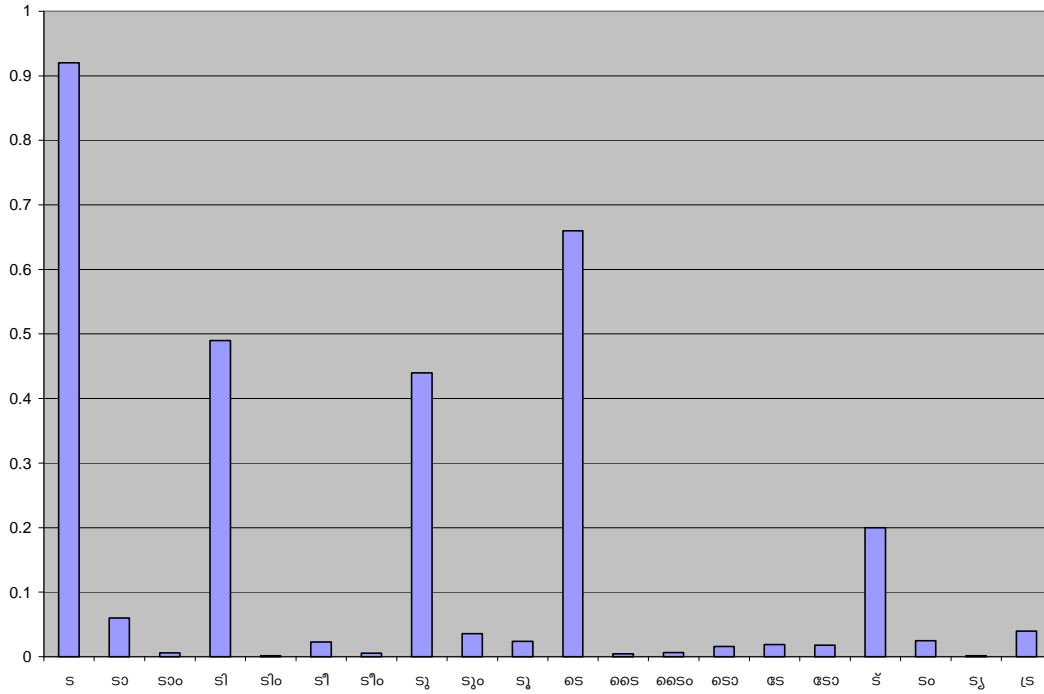
അക്ഷരം	ആവൃത്തി	ആദ്യം	മധ്യം	അന്ത്യം	ശതമാനം
ൺ	183	167	6	1	0.020
ൺാ	358	346	2	0	0.039
ൺി	7	7	0	0	0.0007
ൺു	1	0	0	1	0.00011
ൺ്	154	0	95	0	0.017
ൺെ	18	18	0	0	0.0019
ൺൊ	2	0	2	0	0.000215

ആകെ പരിശോധിച്ച അക്ഷരങ്ങൾ

928971

അക്ഷരം	ആവൃത്തി	ആദ്യം	മധ്യം	അന്ത്യം	ശതമാനം
ട	8586	61	3530	255	0.92
ടാ	559	99	189	22	0.060
ടാഠ	55	3	1	49	0.0059
ടി	4537	798	1123	995	0.49
ടിഠ	12	4	0	0	0.0013
ടിീ	214	155	8	7	0.023
ടിഠ	49	48	0	1	0.0053
ടു	4085	56	1363	172	0.44
ടുഠ	334	0	47	153	0.036
ടു്	227	152	46	25	0.024
ടെ	6151	521	142	4943	0.66
ടെഠ	42	32	4	0	0.0045
ടെഠഠ	62	47	1	14	0.0066
ടെഠാ	146	2	12	0	0.016
ടെഠഠ	17	14	0	1	0.0018
ടെഠ്	154	152	0	0	0.017
ടെഠി	183	46	32	5	0.019
ടെഠിഠ	170	99	44	11	0.018
ടെഠിഠഠ	15	15	0	0	0.0016
ടെഠിഠാ	4	2	0	2	0.0004
ടെഠിഠി	12	12	0	0	0.0013
ടെഠിഠ്	35	33	0	0	0.0038
ടെഠിഠ്	1845	0	119	1545	0.20
ടെഠിഠഠ	236	5	6	220	0.025
ടെഠിഠി	15	0	0	3	0.0016
ടെഠിഠിഠ	13	0	2	1	0.0013
ടെഠിഠിഠി	35	30	2	0	0.0037
ടെഠിഠിഠിഠ	6	0	0	6	0.0007

ശ	375	112	75	50	0.040
ശo	180	120	24	11	0.02
ശീ	262	21	42	60	0.03
ശീ	186	11	40	1	0.02



ആകെ പരിശോധിച്ച അക്ഷരങ്ങൾ 928971

അക്ഷരം	ആവൃത്തി	ആദ്യം	മധ്യം	അന്ത്യം	ശതമാനം
o	308	12	95	14	0.03
oo	16	0	6	0	0.0017
o	275	0	119	6	0.03
oo	17	0	11	6	0.0018
o	11	0	1	1	0.0012
o	13	0	0	8	0.0014
oo	47	0	0	47	0.005
o	9	0	0	2	0.0001
oo	2	0	0	2	0.0009

അക്ഷരം	ആവൃത്തി	ആദ്യം	മധ്യം	അന്ത്യം	ശതമാനം
ഡ	1956	444	426	36	0.210
ഡാ	201	72	62	8	0.022
ഡാം	13	6	0	7	0.0014
ഡി	2284	1119	529	145	0.25
ഡിം	68	0	7	0	0.007
ഡീ	129	62	24	3	0.014
ഡൂ	156	4	60	39	0.017

ഡൂം	42	9	10	21	0.0045
ഡൂ	13	2	6	0	0.0014
ഡൈ	98	81	7	9	0.011
ഡൈഡ	28	10	5	0	0.003
ഡൊ	17	15	0	2	0.0018
ഡൊം	4	4	0	0	0.0008
ഡൊഡ	3	0	1	0	0.0004
ഡേ	157	75	30	43	0.017
ഡോ	521	443	27	15	0.056
ഡൗ	14	8	1	0	0.0015
ഡ്	1848	5	145	1456	0.20
ഡം	40	0	3	31	0.004
ഡ്യ	28	0	12	5	0.003
ഡ്യാ	4	0	4	0	0.0008
ഡ്യൂ	6	2	0	0	0.00075
ഡ്യൂ	44	20	10	0	0.005
ഡ്യാ	125	0	19	98	0.013
ഡ്യാ	33	1	20	2	0.0035
ഡൈ	98	81	7	9	0.0105
ഡൈഡ	28	10	5	0	0.0030
ഡൊ	17	15	0	2	0.0018
ഡൊം	4	4	0	0	0.0008
ഡൊഡ	3	0	1	0	0.00037
ഡേ	157	75	30	43	0.017
ഡോ	521	443	27	15	0.056
ഡ്ര	29	10	16	0	0.0031
ഡ്രാ	31	29	2	0	0.0033
ഡ്രാം	3	3	0	0	0.0003
ഡ്രീ	13	3	9	0	0.0013
ഡ്രിം	2	2	0	0	0.0009

ആകെ പരിശോധിച്ച അക്ഷരങ്ങൾ 928971

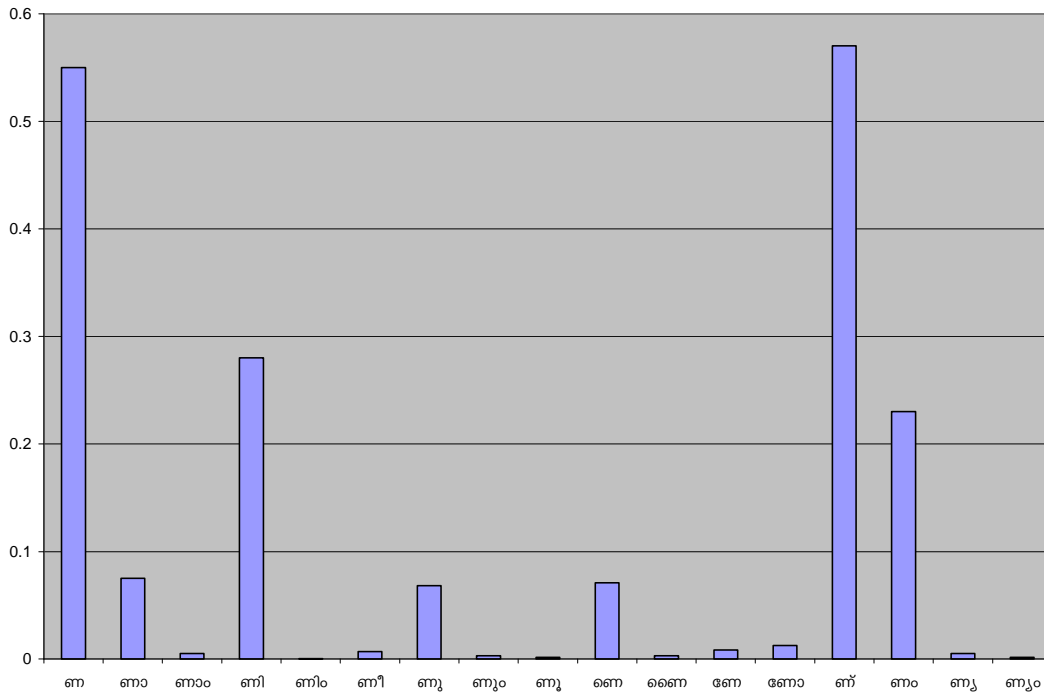
അക്ഷരം	ആവൃത്തി	ആദ്യം	മധ്യം	അന്ത്യം	ശതമാനം
ഡ	17	0	6	5	0.0012
ഡാ	5	0	1	0	0.0005
ഡി	3	0	0	0	0.00032
ഡു	8	0	5	0	0.001
ഡോ	2	0	2	0	0.0002
ഡ്	3	0	2	1	0.0003

ആകെ പരിശോധിച്ച അക്ഷരങ്ങൾ 928971

അക്ഷരം	ആവൃത്തി	ആദ്യം	മധ്യം	അന്ത്യം	ശതമാനം
ണ	5095	6	1335	907	0.55
ണാ	695	0	467	16	0.075
ണാം	47	0	0	47	0.0051
ണി	2582	2	920	457	0.28
ണിം	2	0	1	0	0.0002
ണീ	69	0	22	14	0.007
ണു	635	0	170	310	0.0683552
ണും	26	0	0	26	0.0028
ണൂ	15	0	3	7	0.0016



ണ	664	0	45	18	0.071
ണെ	26	0	15	3	0.0028
ണേ	80	0	53	11	0.0086
ണോ	117	0	22	78	0.0126
ണ്	5321	1	89	5047	0.57
ണം	2101	0	10	2081	0.23
ണ്യ	50	0	9	8	0.005
ണ്യം	13	0	0	13	0.0014



ആകെ പരിശോധിച്ച അക്ഷരങ്ങൾ 928971

അക്ഷരം	ആവൃത്തി	ആദ്യം	മധ്യം	അന്ത്യം	ശതമാനം
ത	9807	2922	2496	1365	1.056
താ	3464	968	780	174	0.37
താം	60	0	4	43	0.006
തി	10289	2110	2797	1701	1.108
തീ	1112	776	122	15	0.12
തു	3723	1334	672	904	0.40
തും	316	0	0	315	0.034
തൂ	96	69	16	0	0.010
തെ	1521	442	116	471	0.16
ത്രെ	27	2	0	25	0.0029
തൈ	42	34	4	2	0.0045
തൊ	507	379	76	0	0.054
തേ	399	168	54	93	0.043
തോ	882	547	94	40	0.095
തൂ	754	237	315	0	0.08

ൽ	4350	4	141	3953	0.47
തം	526	12	71	429	0.057
ത്യ	763	140	207	111	0.082
ത്യം	134	6	50	3	0.014
ത്യു	56	0	6	38	0.006
ത്യൂ	22	0	13	9	0.002
ത്യും	76	0	0	74	0.008
ത	309	17	146	35	0.033
തം	144	0	0	143	0.015
ത്ര	1676	5	711	292	0.18
ത്രം	79	1	24	21	0.009
ത്രീ	463	45	85	194	0.049
ത്രീ	238	6	88	36	0.026
ത്രൂ	19	0	7	2	0.0020
ത്രും	497	0	5	491	0.054

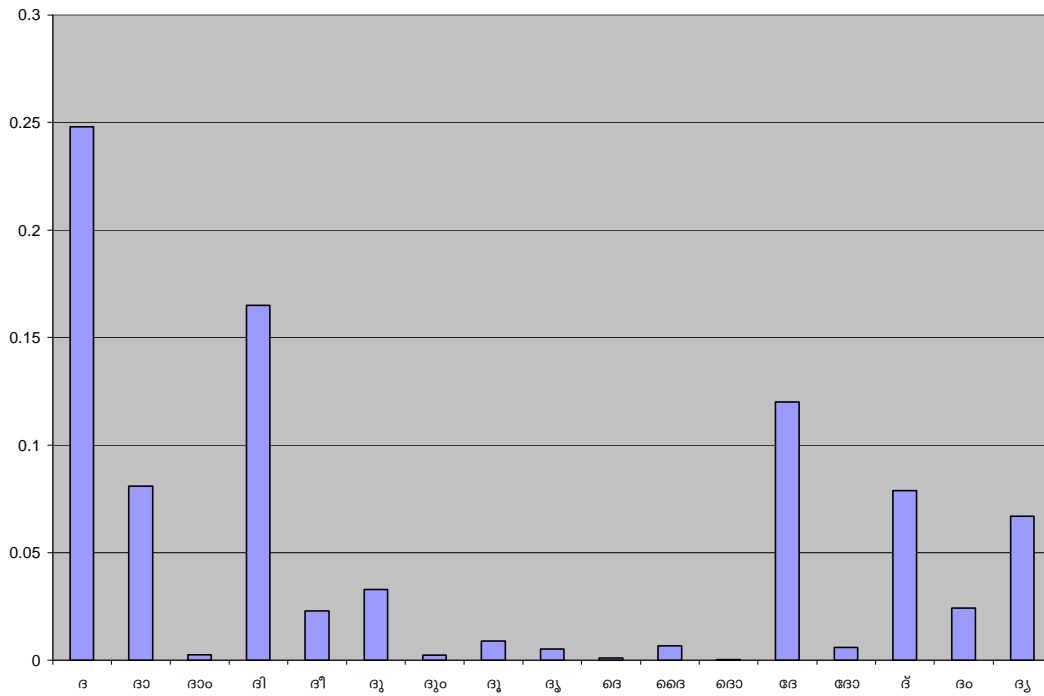
ആകെ പരിശോധിച്ച അക്ഷരങ്ങൾ

928971

അക്ഷരം	ആവൃത്തി	ആദ്യം	മധ്യം	അന്ത്യം	ശതമാനം
ഥ	370	0	126	100	0.04
ഥം	210	7	97	3	0.023
ഥി	230	0	86	88	0.027
ഥീ	3	3	0	0	0.00037
ഥു	4	0	4	0	0.00083
ഥൂ	1	1	0	0	0.00045
ഥേ	23	0	22	0	0.0025
ഥ്	36	0	0	36	0.0039
ഥം	25	0	0	25	0.0027
ഥ്യ	13	0	3	1	0.0014

അക്ഷരം	ആവൃത്തി	ആദ്യം	മധ്യം	അന്ത്യം	ശതമാനം
ദ	2303	306	684	206	0.248
ദം	753	152	205	29	0.081
ദാം	23	0	15	2	0.0025
ദി	1537	554	345	142	0.165
ദീ	216	107	53	9	0.023
ദു	309	185	19	8	0.033
ദും	21	4	0	17	0.0023
ദൂ	85	52	5	0	0.009
ദൂ	49	23	7	0	0.0053
ദെ	9	7	0	0	0.00097
ദൈ	63	53	4	0	0.0068
ദൊ	3	1	1	0	0.00032
ദേ	1117	387	353	1	0.12
ദോ	60	47	2	1	0.006
ദേ്യ	4	0	0	0	0.0004
ദേ്യം	162	3	35	0	0.0174
ദേ്യ	11	0	3	0	0.00118
ദഃ	30	24	1	0	0.0032
ദ്	734	57	32	304	0.079
ദ്	220	10	6	191	0.0243

ഭൂ	628	1	242	162	0.067
ഭൂറ	541	0	220	13	0.058
ഭൂരൂ	109	0	77	0	0.012
ഭൂരൂരൂ	20	0	20	0	0.0021
ഭൂരൂരൂരൂ	2	0	0	2	0.00022
ഭൂരൂരൂരൂരൂ	148	0	4	138	0.016
ഭൂരൂരൂരൂരൂരൂ	8	2	2	0	0.00086
ഭൂരൂരൂരൂരൂരൂരൂ	12	6	6	0	0.0013
ഭൂരൂരൂരൂരൂരൂരൂരൂ	17	11	2	0	0.0018
ഭൂരൂരൂരൂരൂരൂരൂരൂരൂ	13	0	7	0	0.0014
ഭൂരൂരൂരൂരൂരൂരൂരൂരൂരൂ	131	16	45	11	0.014
ഭൂരൂരൂരൂരൂരൂരൂരൂരൂരൂരൂ	38	5	20	2	0.0041
ഭൂരൂരൂരൂരൂരൂരൂരൂരൂരൂരൂരൂ	10	0	0	10	0.0011
ഭൂരൂരൂരൂരൂരൂരൂരൂരൂരൂരൂരൂരൂ	2	0	0	2	0.0002

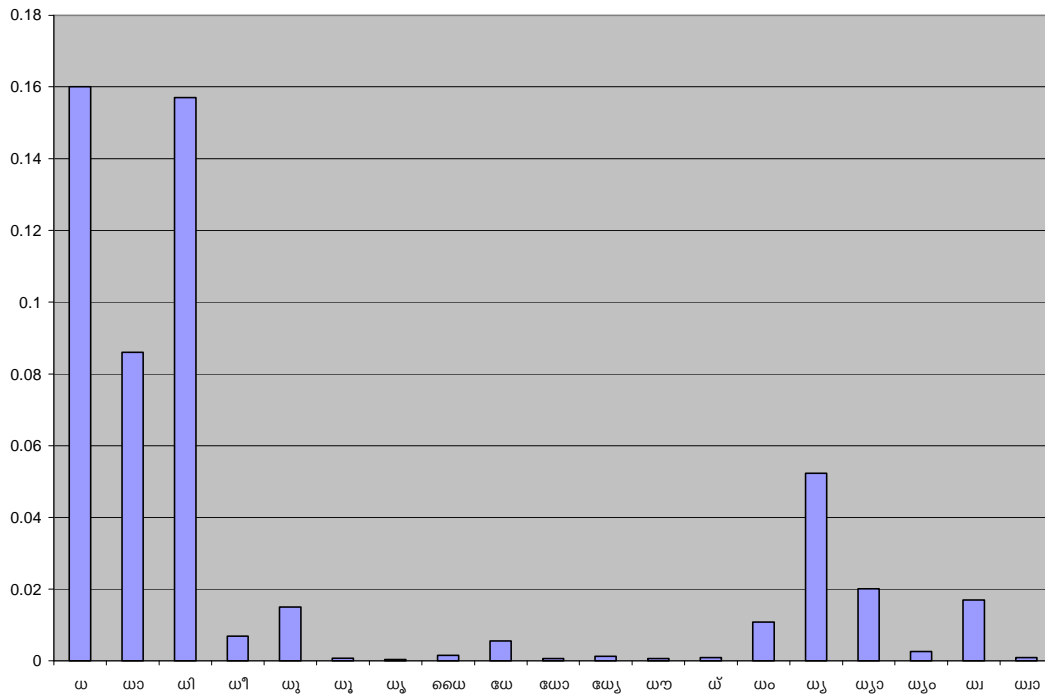


ആകെ പരിശോധിച്ച അക്ഷരങ്ങൾ

928971

അക്ഷരം	ആവൃത്തി	ആദ്യം	മധ്യം	അന്ത്യം	ശതമാനം
ധ	1500	242	355	282	0.16
ധാ	802	82	271	7	0.086
ധി	1463	6	424	280	0.157
ധീ	64	6	37	0	0.0069
ധൂ	137	4	75	14	0.015
ധൂര	7	7	0	0	0.00075
ധൂരൂ	3	3	0	0	0.00032
ധൈ	14	14	0	0	0.0015

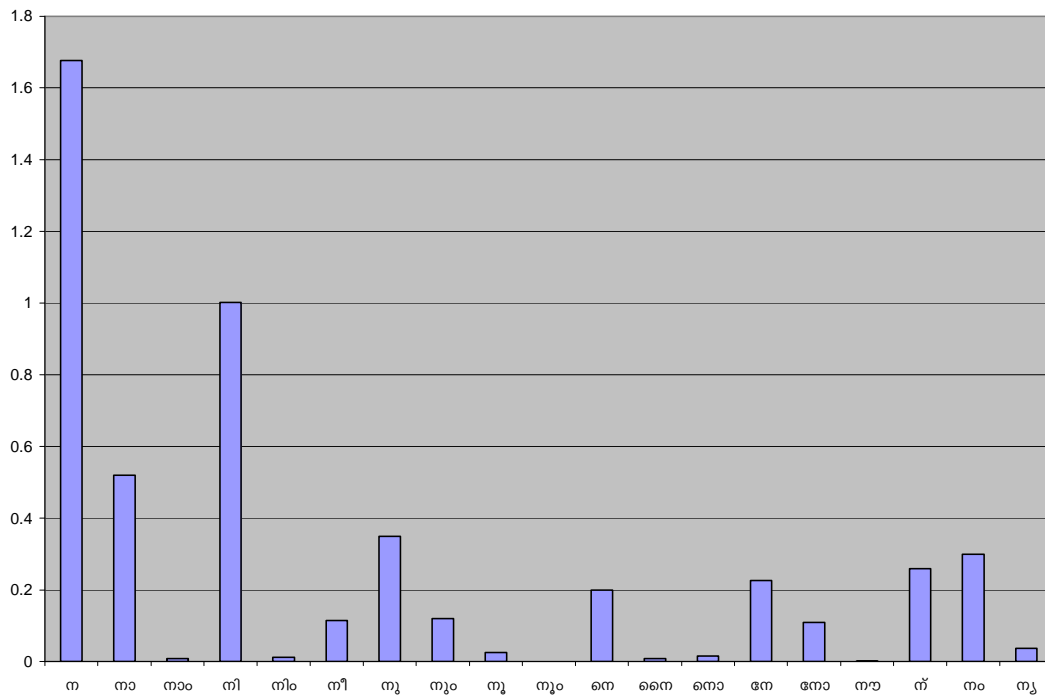
ഘ	52	0	18	0	0.0056
ഘാ	6	0	6	0	0.00065
ഘ്യ	11	0	2	6	0.0012
ഘൗ	6	0	0	6	0.0006
ഘ്	8	0	0	1	0.00086
ഘം	100	0	0	100	0.01076
ഘ്യം	486	0	210	24	0.0523
ഘ്യൗ	187	6	65	3	0.0201
ഘ്യം	24	0	2	22	0.0026
ഘി	162	116	13	0	0.017
ഘാ	8	1	5	0	0.00086



ആകെ പരിശോധിച്ച അക്ഷരങ്ങൾ 928971

അക്ഷരം	ആവൃത്തി	ആദ്യം	മധ്യം	അന്ത്യം	ശതമാനം
ന	15569	4801	4078	2064	1.68
നാ	4827	1532	730	140	0.52
നാം	76	62	2	4	0.0081
നി	9312	4713	1697	592	1.002
നിം	111	0	23	1	0.0119
നീ	1071	456	388	4	0.1153
നു	3238	104	733	903	0.349
നും	1116	45	2	1067	0.120
നു്	238	109	30	0	0.026
നും	2	0	0	2	0.00022
നെ	1874	342	203	839	0.20
നൈ	77	43	12	1	0.008
നൊ	137	21	23	1	0.015
നേ	2104	1438	268	31	0.226

ഭനാ	978	217	327	62	0.11
ഭന	217	0	130	0	0.023
ന	31	27	0	0	0.0033
ന	9	0	3	2	0.00097
ന	2409	489	28	1849	0.259
ന	2820	58	70	2656	0.30
ന	347	2	122	128	0.037
ന	116	30	25	5	0.012
ന	210	178	1	28	0.023
ന	149	0	2	147	0.016
ന	21	0	14	0	0.002

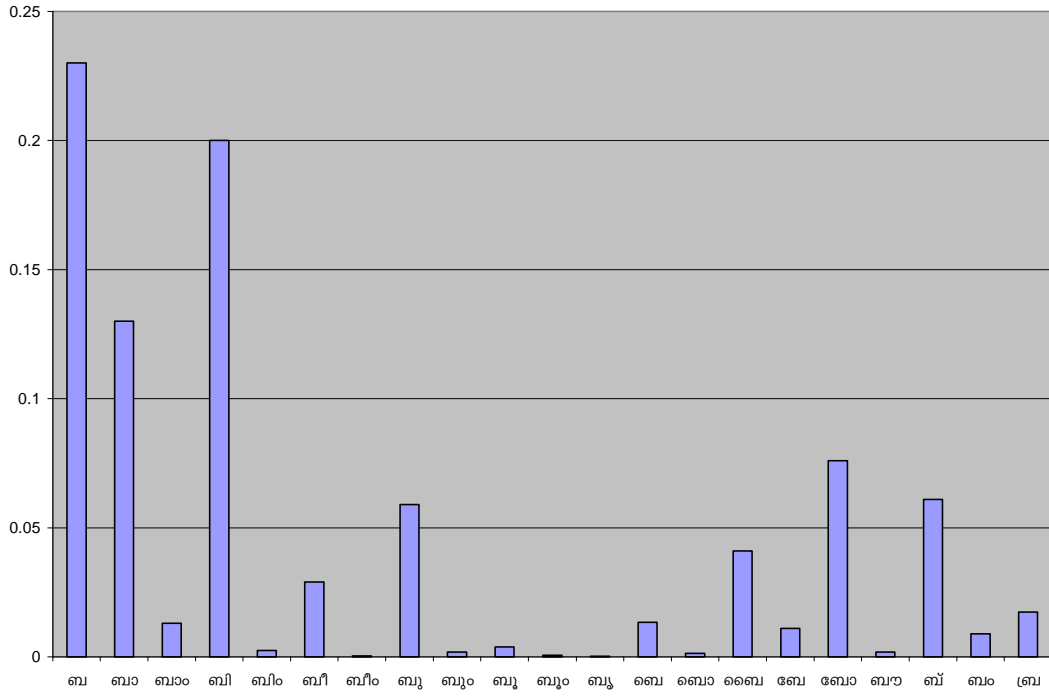


അക്ഷരം	ആവൃത്തി	ആദ്യം	മധ്യം	അന്ത്യം	ശതമാനം
വ	13331	8152	1761	506	1.44
പാ	3110	1802	528	14	0.335
പാഠം	9	5	4	0	0.0009
പി	3868	2870	383	56	0.416
പിഠം	1	1	0	0	0.00011
പീ	178	88	48	1	0.019
പൂ	3057	1732	67	20	0.33
പൂഠം	8	0	1	7	0.00086
പൂ	729	515	97	1	0.078
പെ	871	611	140	5	0.094
പൈ	101	95	1	0	0.011
പൊ	647	613	12	0	0.070
പോ	2064	1332	164	18	0.22
പ്ര	5408	4683	359	6	0.58
പൃ	13	12	0	0	0.0014
പൗ	80	65	8	0	0.0086
പ്	839	33	283	153	0.090
പഠം	296	2	0	286	0.032
പ്യ	103	1	23	4	0.011
പ്യഠ	11	11	0	0	0.0012
പ്യൂ	2	0	2	0	0.00021
പ്യൂ	102	8	74	0	0.011
പ്രാ	578	342	110	2	0.06
പ്രാഠം	3	0	1	2	0.0003
പ്രി	170	136	22	0	0.0183
പ്രിഠം	8	0	0	8	0.00086
പ്രീ	81	75	1	3	0.0087
പ്രീഠം	67	0	0	45	0.0072
പ്രൊ	135	135	0	0	0.0145

അക്ഷരം	ആവൃത്തി	ആദ്യം	മധ്യം	അന്ത്യം	ശതമാനം
ഫ	791	312	174	66	0.085
ഫാ	267	221	25	8	0.029
ഫാഓ	14	8	0	6	0.0015
ഫി	681	306	153	37	0.073
ഫിഓ	8	1	7	0	0.00086
ഫീ	537	167	243	0	0.0578
ഫു	176	143	20	2	0.019
ഫുഓ	15	6	0	9	0.0016
ഫു	2	0	2	0	0.00021
ഫെ	171	124	21	0	0.018
ഫൈ	163	105	42	1	0.018
ഫൊ	11	11	0	0	0.0012
ഫേ	30	15	2	4	0.003
ഫോ	991	778	75	1	0.11
ഫോഓ	26	13	0	13	0.0028
ഫൗ	10	10	0	0	0.0011
ഫ്	993	59	115	706	0.11
ഫ്യ	1	1	0	0	0.00011
ഫ്യൂ	2	0	2	0	0.00022
ഫ്ര	49	40	1	0	0.005
ഫ്രാ	40	35	0	0	0.0043
ഫ്രി	77	2	21	3	0.0083
ഫ്രീ	53	51	0	1	0.00570

അക്ഷരം	ആവൃത്തി	ആദൃം	മധ്യം	അന്ത്യം	ശതമാനം
ബ	2182	1011	467	45	0.23
ബാ	1242	890	119	1	0.13
ബാം	119	49	52	0	0.013
ബി	1867	1358	189	111	0.20
ബിം	22	9	10	0	0.0024
ബീ	267	172	44	0	0.029
ബീം	3	3	0	0	0.000323
ബു	549	254	104	120	0.059
ബും	17	10	0	7	0.0018
ബൂ	36	24	5	1	0.0039
ബൂം	6	3	0	3	0.00064
ബ്യ	2	2	0	0	0.00021
ബെ	125	91	9	9	0.0134
ബൊ	13	13	0	0	0.0014
ബൈ	380	113	138	70	0.041
ബേ	104	71	23	0	0.011
ബോ	704	473	53	12	0.076
ബോം	76	74	0	0	0.0082
ബൗ	18	13	4	0	0.0019
ബ്	570	3	121	274	0.061
ബം	84	47	9	18	0.0090
ബ്യ	28	0	6	11	0.0030
ബ്യൂ	2	1	1	0	0.000215
ബ്യൂ	55	37	9	1	0.0059
ബ്ര	162	93	36	6	0.0174
ബ്രൊ	97	76	14	0	0.0104
ബ്രാം	2	0	2	0	0.000215
ബ്രി	117	80	22	0	0.0126



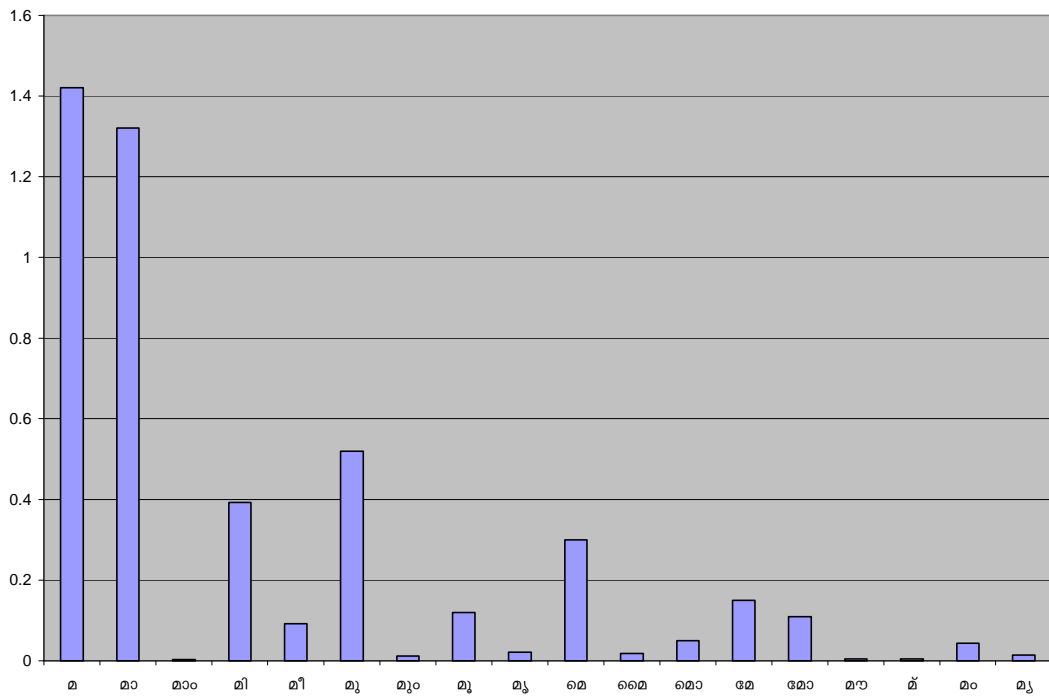


ആകെ പരിശോധിച്ച അക്ഷരങ്ങൾ

928971

അക്ഷരം	ആവൃത്തി	ആദൃം	മധ്യം	അന്ത്യം	ശതമാനം
ഭ	1909	790	386	130	0.205
ഭാ	2168	1174	320	65	0.233
ഭാഠ	11	0	4	0	0.0011
ഭി	1350	24	598	27	0.145
ഭീ	130	111	10	5	0.014
ഭൂ	57	21	13	4	0.0061
ഭൂഠ	524	209	23	0	0.056
ഭേ	57	36	6	0	0.0061
ഭോ	58	22	15	3	0.0062
ഭ്യ	3	3	0	0	0.00032
ഭ്യഠ	1	1	0	0	0.00011
ഭൗ	21	15	1	0	0.002
ഭ്	8	0	6	0	0.00086
ഭം	107	36	6	56	0.0115
ഭ്യ	190	0	90	1	0.020
ഭ്യഠ	178	0	31	3	0.0191
ഭ്യു	2	0	0	0	0.000215
ഭ്യുഠ	2	0	0	0	0.000215
ഭ്യം	10	0	0	10	0.0011
ഭ്ര	19	10	2	0	0.0020
ഭ്രഠ	4	2	1	0	0.0004

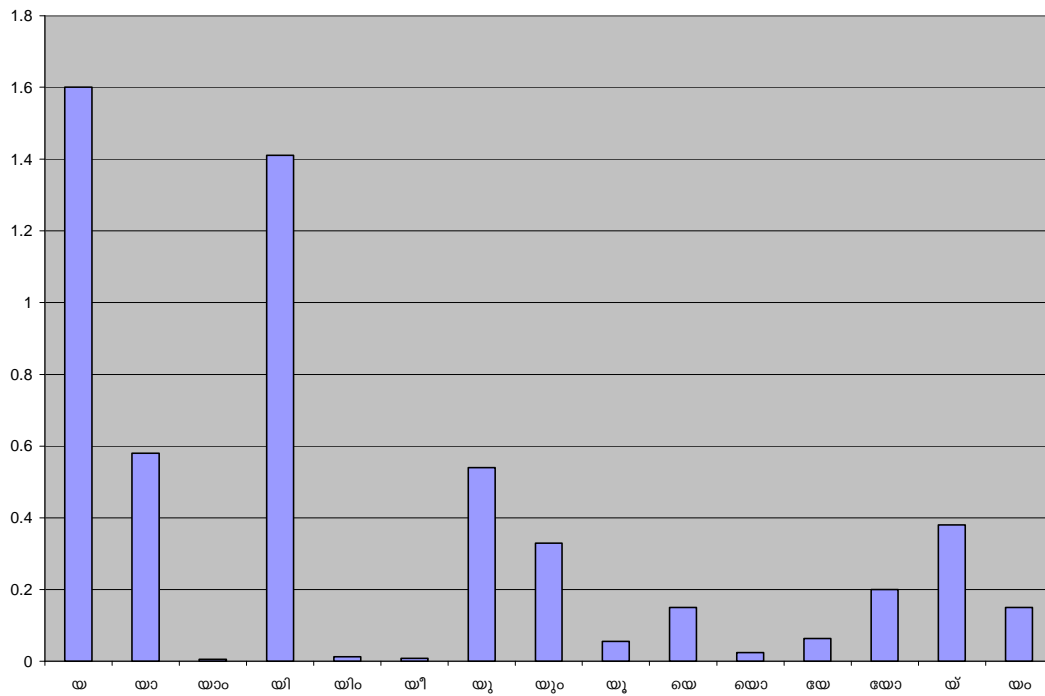
അക്ഷരം	ആവൃത്തി	ആദൃം	മധ്യം	അന്ത്യം	ശതമാനം
മ	13154	6037	2616	634	1.42
മാ	12294	3116	1824	48	1.32
മാം	29	12	9	3	0.0031
മി	3651	684	1040	571	0.393
മിം	14	0	7	0	0.0015
മീ	858	383	288	4	0.092
മൂ	4832	2592	861	5	0.52
മൂം	108	97	0	11	0.0116
മു	1132	728	106	1	0.12
മ്യ	193	122	28	6	0.021
മെ	2816	696	207	53	0.30
മൈ	165	154	2	2	0.018
മൊ	480	263	78	0	0.05
മേ	1427	578	279	178	0.15
മോ	996	516	145	144	0.11
മൗ	42	39	2	0	0.0045
മ്	45	5	15	15	0.0048
മം	408	43	50	279	0.044
മ്യ	128	36	17	46	0.0138
മ്യം	12	4	2	3	0.0014
മ്യൂ	6	5	0	0	0.00065
മ്യൂ	68	44	10	0	0.0073
മ്യം	14	0	0	14	0.0015



ആകെ പരിശോധിച്ച അക്ഷരങ്ങൾ

928971

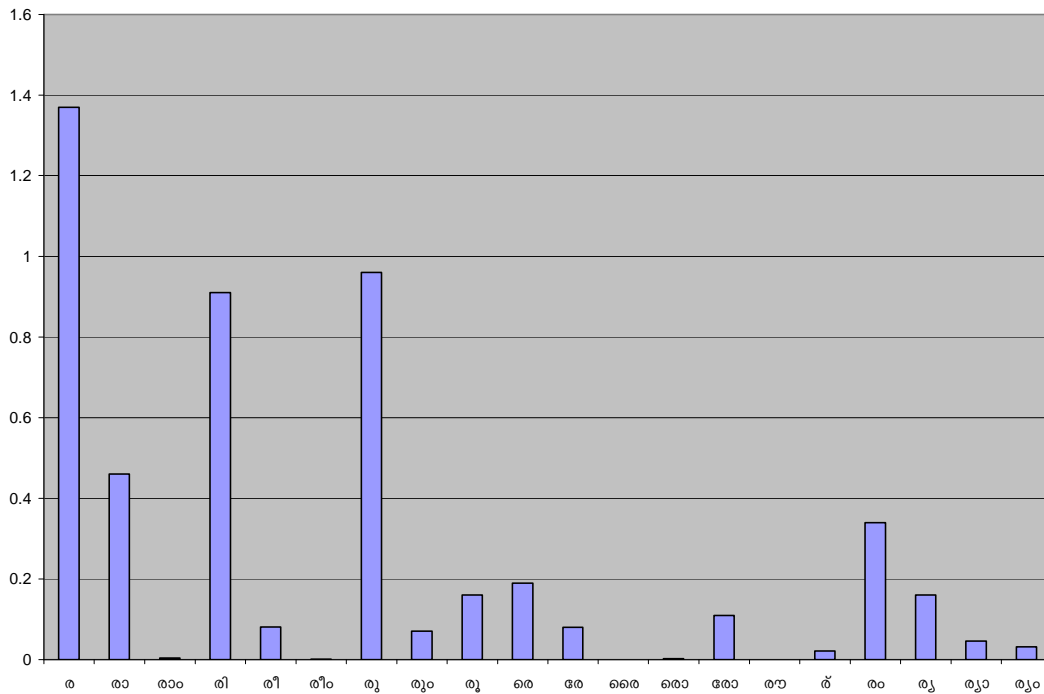
അക്ഷരം	ആവൃത്തി	ആദ്യം	മധ്യം	അന്ത്യം	ശതമാനം
യ	14891	156	3529	4909	1.60
യാ	5390	307	1047	52	0.580
യാം	48	0	4	42	0.0052
യി	13070	83	2351	3534	1.41
യിം	112	0	71	24	0.012
യീ	77	1	8	8	0.008
യു	4986	678	875	20	0.54
യും	3043	36	3	2997	0.33
യൂ	517	353	42	28	0.056
യെ	1360	37	233	553	0.15
യൊ	225	0	47	7	0.024
യേ	595	14	148	54	0.064
യോ	1851	506	354	290	0.20
യ്	3522	25	1297	259	0.38
യം	1423	2	17	1384	0.15



ആകെ പരിശോധിച്ച അക്ഷരങ്ങൾ

928971

അക്ഷരം	ആവൃത്തി	ആദ്യം	മധ്യം	അന്ത്യം	ശതമാനം
ര	12722	1774	3731	947	1.37
രാ	4262	1607	973	59	0.46
രാം	33	3	9	19	0.0036
രി	8487	5	2790	920	0.91
രീ	751	146	282	5	0.081
രീം	7	0	0	7	0.0007
രു	8952	83	1402	1589	0.96
രും	658	0	10	640	0.071
രു	1458	1202	77	34	0.16
രെ	1734	1	154	1342	0.19
രേ	747	122	165	87	0.08
രൈ	3	0	0	2	0.0003
രൊ	25	6	6	4	0.0027
രോ	1019	190	249	157	0.11
രൗ	2	2	0	0	0.0002
ർ	203	1	44	124	0.022
രം	3177	155	147	2584	0.34
ര്യ	1478	1	505	535	0.16
ര്യം	426	0	42	4	0.046
ര്യം	296	0	0	296	0.032

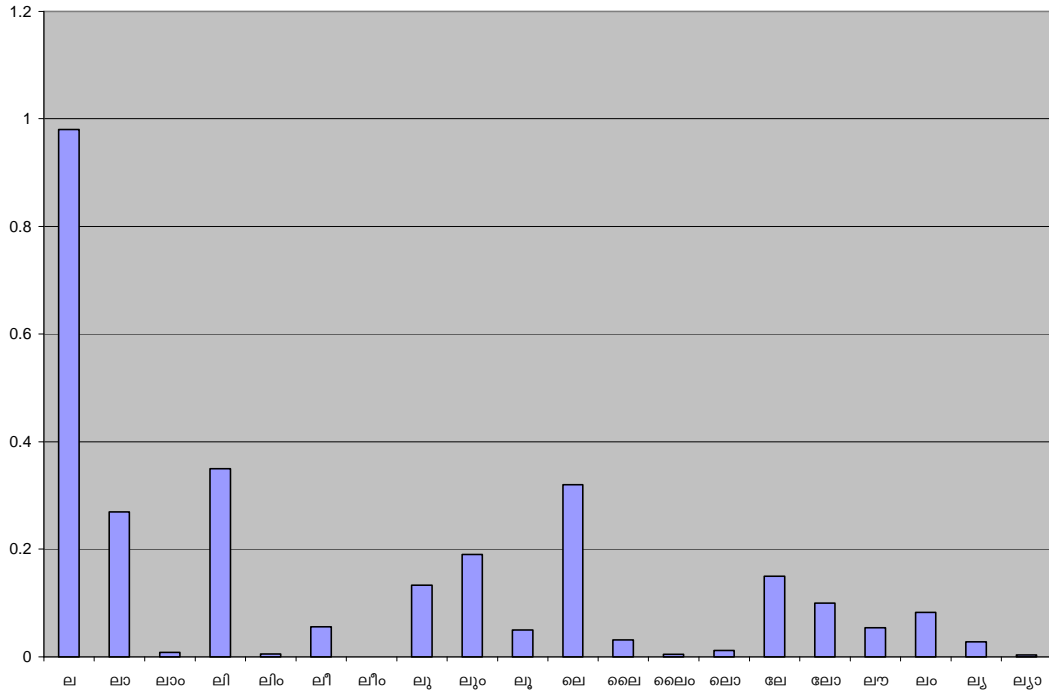


ആകെ പരിശോധിച്ച അക്ഷരങ്ങൾ

928971

അക്ഷരം	ആവൃത്തി	ആദ്യം	മധ്യം	അന്ത്യം	ശതമാനം
റ	8224	425	3460	295	0.89
റാ	1168	206	468	28	0.13
റാം	46	4	0	42	0.005
റി	5185	635	1291	1050	0.56
റിം	177	2	10	0	0.020
റീ	169	83	26	0	0.018
റീം	2	2	0	0	0.0002
റൂ	1519	14	505	156	0.16
റൂം	255	2	3	248	0.027
റൂ	103	43	36	4	0.011
റൂം	29	18	0	11	0.0031
റെ	543	194	52	223	0.058
റെ	26	9	6	0	0.0028
റൊ	31	23	4	0	0.0033
രേ	486	54	210	23	0.052
രോ	705	406	123	64	0.076
രോം	6	2	0	4	0.00065
രൂ	2	2	0	0	0.0002
രൗ	80	46	23	3	0.0086
ര്	497	0	36	423	0.054
റം	222	9	14	180	0.024

അക്ഷരം	ആവൃത്തി	ആദ്യം	മധ്യം	അന്ത്യം	ശതമാനം
ല	9077	1399	2601	1401	0.98
ലാ	2511	222	632	95	0.27
ലാം	77	8	14	45	0.0083
ലി	3242	217	1058	510	0.35
ലിം	48	0	13	27	0.0052
ലീ	520	163	282	4	0.056
ലീം	1	0	0	1	0.00011
ലു	1244	8	347	89	0.133
ലും	1766	3	16	1733	0.19
ലു	483	18	78	0	0.05
ലൈ	2955	121	112	2366	0.32
ലൈ	298	174	38	18	0.032
ലൈം	47	47	0	0	0.005
ലൊ	110	7	22	1	0.012
ലേ	1366	198	193	61	0.15
ലോ	949	425	122	191	0.10
ലോം	10	5	0	3	0.0011
ലൗ	12	7	0	0	0.0013
ല്	503	1	169	88	0.054
ലം	770	30	29	658	0.083
ല്യ	263	2	71	65	0.028
ല്യാ	31	0	14	1	0.0033
ല്യു	3	0	3	0	0.0003
ല്യൂ	2	0	1	0	0.0002
ല്യൂം	73	0	9	63	0.0079
ല്യാ	2	0	2	0	0.0002



ആകെ പരിശോധിച്ച അക്ഷരങ്ങൾ

928971

അക്ഷരം	ആവൃത്തി	ആദ്യം	മധ്യം	അന്ത്യം	ശതമാനം
ഉ	6311	3	1448	1707	0.68
ഉറ	1007	2	231	50	0.11
ഉറം	6	0	0	6	0.00065
ഉി	4263	1	882	386	0.46
ഉിം	12	0	2	0	0.0013
ഉീ	87	0	49	0	0.0094
ഉു	1814	0	188	41	0.20
ഉും	825	0	0	820	0.089
ഉൂ	65	0	5	41	0.0070
ഉെ	1015	0	55	653	0.109
ഉൈ	10	0	0	1	0.00108
ഉൊ	66	0	8	0	0.00710
ഉേ	369	0	227	10	0.039
ഉോ	258	0	57	30	0.028
ഉ്	46	0	3	25	0.0050
ഉം	910	0	31	855	0.098

ആകെ പരിശോധിച്ച അക്ഷരങ്ങൾ

928971

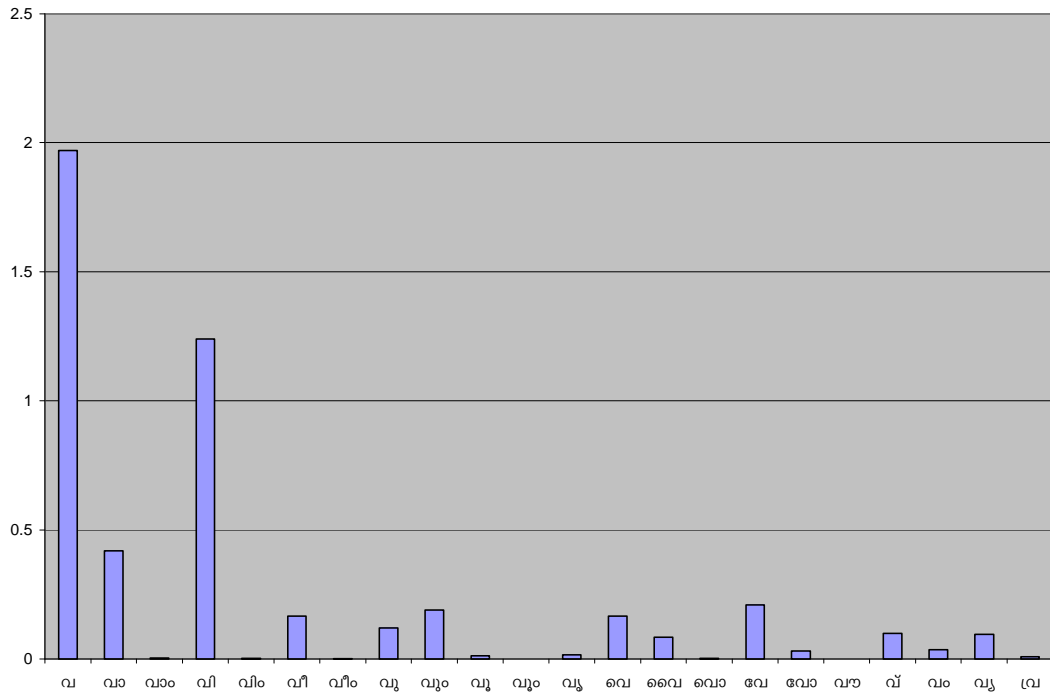
അക്ഷരം	ആവൃത്തി	ആദ്യം	മധ്യം	അന്ത്യം	ശതമാനം
ഴ	1056	0	380	230	0.114
ഴറ	137	0	53	0	0.015
ഴറം	16	0	0	16	0.0017
ഴി	2372	0	1235	207	0.26
ഴീ	10	0	9	0	0.00108

ശു	709	0	299	22	0.076
ശും	111	0	0	111	0.012
ശു	16	0	16	0	0.0017
ശെ	112	0	15	56	0.0120
ശൊ	10	0	1	0	0.00108
ശേ	35	0	4	12	0.00377
ശോ	10	0	7	3	0.00108
ശ്	1946	0	559	132	0.209
ശം	34	0	2	31	0.0037

ആകെ പരിശോധിച്ച അക്ഷരങ്ങൾ

928971

അക്ഷരം	ആവൃത്തി	ആദ്യം	മധ്യം	അന്ത്യം	ശതമാനം
വ	18328	4932	4404	613	1.97
വാ	3900	1773	757	33	0.42
വാം	34	0	0	34	0.0037
വീ	11555	6708	2037	464	1.24
വീം	19	8	7	0	0.0020
വീ	1554	1079	195	7	0.167
വീം	6	0	3	0	0.00065
വു	1104	16	360	101	0.12
വും	1746	3	5	1728	0.19
വു	111	8	22	16	0.012
വും	1	0	0	1	0.000108
വൂ	149	58	33	0	0.016
വെ	1555	1015	111	33	0.167
വൈ	785	668	38	0	0.085
വൊ	21	12	4	0	0.0023
വേ	1995	950	420	104	0.21
വോ	297	168	42	38	0.032
വൂ	2	2	0	0	0.00022
വൗ	3	3	0	0	0.00032
വ്	922	8	28	846	0.099
വം	341	24	27	258	0.037
വ്യ	894	662	84	30	0.096
വ്യം	356	287	18	0	0.038
വ്യു	11	4	0	7	0.0012
വ്യു	52	11	2	28	0.0056
വ്യും	71	28	0	43	0.0076
വ്ര	82	5	19	7	0.0088



ആകെ പരിശോധിച്ച അക്ഷരങ്ങൾ

928971

അക്ഷരം	ആവൃത്തി	ആദ്യം	മധ്യം	അന്ത്യം	ശതമാനം
ശ	3205	1187	583	194	0.345
ശാ	982	361	232	1	0.106
ശി	984	227	263	85	0.106
ശിം	9	9	0	0	0.00097
ശീ	458	45	269	6	0.049
ശു	419	158	93	15	0.045
ശും	4	0	0	4	0.00043
ശൂ	141	13	111	0	0.015
ശൗ	19	12	3	0	0.0020
ശൌ	27	19	3	0	0.0029
ശോ	15	11	1	0	0.0016
ശേ	718	269	143	3	0.077
ശോ	303	79	132	0	0.033
ശേ	8	3	1	3	0.00086
ശൗ	2	2	0	0	0.000215
ശ്	441	13	94	109	0.047
ശം	332	6	21	276	0.036
ശ്യ	681	1	464	11	0.073
ശ്യാ	24	20	0	1	0.0026
ശ്യാം	11	11	0	0	0.0012
ശ്യം	113	2	0	110	0.012



ശ്യാ	299	1	104	1	0.032
ശ്യാഠ	205	8	65	0	0.022
ശ്യാഠഠ	9	0	1	0	0.00097
ശ്യാഠഠഠ	506	315	95	4	0.054
ശ്യാഠഠഠഠ	13	6	7	0	0.0014
ശ്യാഠഠഠഠഠ	22	13	2	0	0.0024
ശ്യാഠഠഠഠഠഠ	506	444	0	54	0.054
ശ്യാഠഠഠഠഠഠഠ	12	12	0	0	0.0013
ശ്യാഠഠഠഠഠഠഠഠ	31	3	14	0	0.0033

ആകെ പരിശോധിച്ച അക്ഷരങ്ങൾ

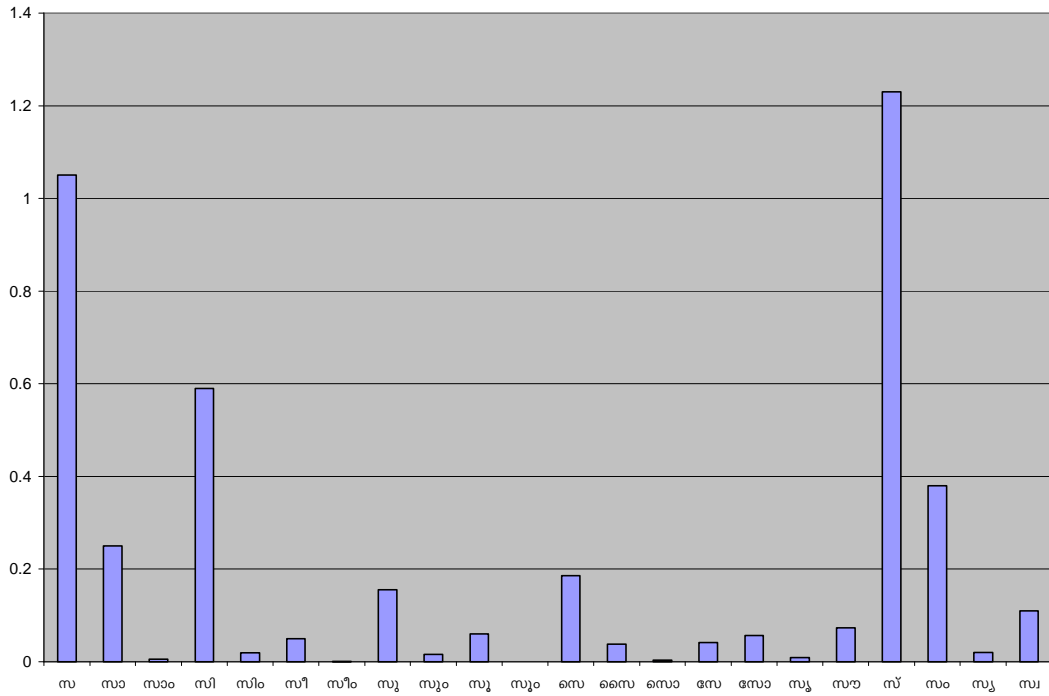
928971

അക്ഷരം	ആവൃത്തി	ആദ്യം	മധ്യം	അന്ത്യം	ശതമാനം
ഷ	4036	88	1182	333	0.43
ഷഠ	204	108	42	21	0.022
ഷഠഠ	26	0	13	3	0.0028
ഷഠഠഠ	796	90	196	72	0.086
ഷഠഠഠഠ	12	1	6	0	0.0013
ഷഠഠഠഠഠ	125	45	57	0	0.013
ഷഠഠഠഠഠഠ	78	4	62	2	0.0084
ഷഠഠഠഠഠഠഠ	13	0	0	13	0.0014
ഷഠഠഠഠഠഠഠഠ	29	18	7	0	0.0031
ഷെ	74	63	6	3	0.0080
ഷൈ	29	29	0	0	0.0031
ഷൈഠ	5	4	0	0	0.00054
ഷേ	168	19	39	13	0.018
ഷേഠ	213	142	36	10	0.023
ഷ്	2107	10	516	484	0.227
ഷ്ഠ	649	9	14	621	0.070
ഷ്യ	466	0	142	45	0.050
ഷ്യഠ	39	0	23	5	0.0042
ഷ്യാ	20	2	14	0	0.0022
ഷ്യാഠ	33	33	0	0	0.0036

ആകെ പരിശോധിച്ച അക്ഷരങ്ങൾ

928971

അക്ഷരം	ആവൃത്തി	ആദ്യം	മധ്യം	അന്ത്യം	ശതമാനം
സ	9766	5741	1581	352	1.05
സാ	2352	1229	417	26	0.25
സാം	52	47	2	1	0.0056
സി	5494	1879	1563	332	0.59
സീ	177	94	12	6	0.019
സീ	465	354	56	0	0.050
സീ	8	0	6	2	0.0008
സു	1442	877	302	8	0.155
സു	145	6	0	139	0.016
സു	556	444	53	3	0.060
സു	2	2	0	0	0.000215
സെ	1728	1599	64	10	0.186
സൈ	353	181	110	0	0.038
സൊ	32	29	2	0	0.0034
സേ	387	179	106	5	0.042
സോ	526	262	31	23	0.057
സൂ	81	59	9	0	0.0087
സൗ	675	621	14	0	0.073
സ്	11380	1270	1451	6601	1.23
സം	3495	2369	246	671	0.38
സ്യ	189	19	44	48	0.020
സ്യ	22	2	20	0	0.0024
സ്യ	25	6	14	0	0.0027
സ്യ	44	0	2	41	0.0047
സ്വ	977	858	36	0	0.11
സ്വാ	428	355	23	0	0.046
സ്വീ	22	17	0	1	0.0024
സ്വീ	183	177	2	0	0.020
സ്വം	45	0	4	41	0.0048
സ്ര	35	5	12	4	0.0038



ആകെ പരിശോധിച്ച അക്ഷരങ്ങൾ

928971

ഹ	2366	476	804	73	0.25
ഹാ	932	272	215	6	0.10
ഹാം	28	2	4	20	0.0030
ഹി	1358	173	559	114	0.15
ഹിം	25	0	9	12	0.003
ഹീ	86	40	33	0	0.009
ഹ്യ	193	51	66	7	0.021
ഹ്യം	25	11	3	0	0.0027
ഹ്യം	1	1	0	0	0.00010
ഹെ	233	223	7	1	0.025
ഹൈ	226	219	4	0	0.024
ഹൈം	1	1	0	0	0.00010
ഹൊ	10	10	0	0	0.00108
ഹേ	86	10	53	5	0.0093
ഹോ	536	381	27	5	0.058
ഹോം	44	43	0	1	0.0047
ഹ്യ	180	56	58	1	0.019
ഹൗ	83	75	0	0	0.0090
ഹ്	58	1	21	14	0.0062
ഹിം	603	42	0	554	0.065

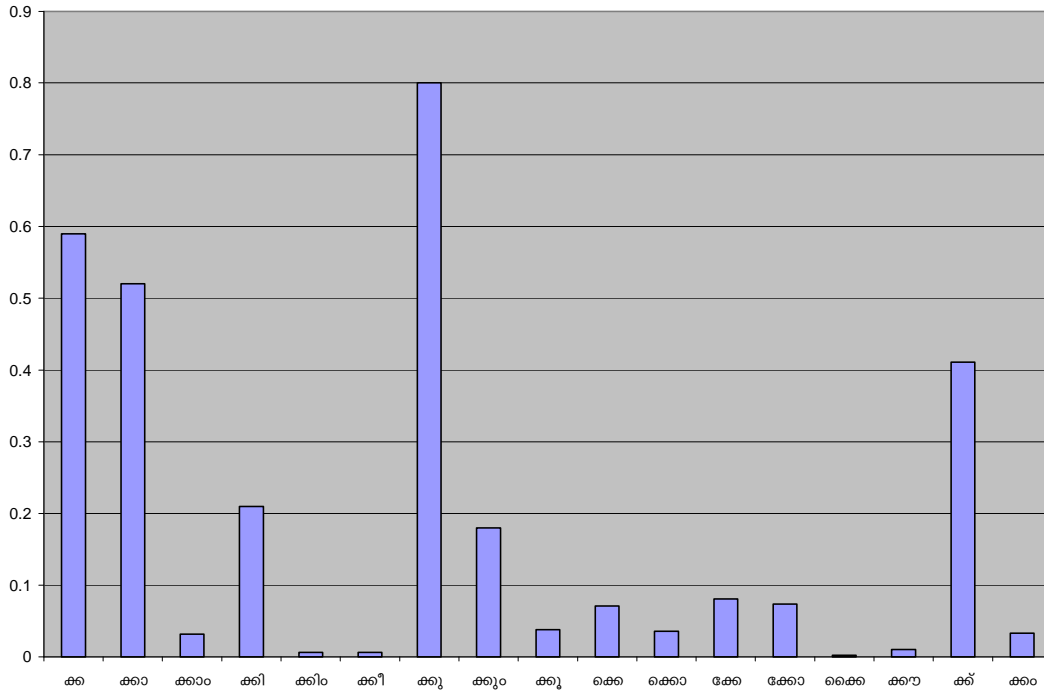
ഹൃ	60	0	28	14	0.0065
ഹൃ	10	10	0	0	0.00108
ഹൃ	23	0	21	0	0.00248

### 3.3.3 കൂട്ടക്ഷരങ്ങൾ

ആകെ പരിശോധിച്ച അക്ഷരങ്ങൾ

928971

അക്ഷരം	ആവൃത്തി	ആദ്യം	മധ്യം	അന്ത്യം	ശതമാനം
ക്ക	5509	6	1929	244	0.59
ക്കൊ	4794	6	1108	16	0.52
ക്കോ	299	1	2	295	0.032
ക്കി	1953	0	483	560	0.21
ക്കിം	58	0	25	21	0.006
ക്കീ	58	0	33	10	0.006
ക്കു	7437	24	1600	828	0.80
ക്കും	1651	3	5	1627	0.18
ക്കു	351	0	83	59	0.038
ക്കും	2	0	0	2	0.0002
ക്കെ	664	2	203	225	0.071
ക്കൊ	334	1	110	2	0.036
ക്കേ	751	1	310	55	0.081
ക്കോ	689	0	181	70	0.074
ക്കൈ	19	0	10	7	0.0020
ക്കു	17	0	0	0	0.0018
ക്കു	94	0	67	0	0.010
ക്കി	3826	24	41	3704	0.411
ക്കിം	306	0	8	286	0.033
ക്കു	2	0	0	0	0.0002



ആകെ പരിശോധിച്ച അക്ഷരങ്ങൾ

928971

അക്ഷരം	ആവൃത്തി	ആദ്യം	മധ്യം	അന്ത്യം	ശതമാനം
ക	1851	83	578	381	0.20
കറ	252	11	46	83	0.027
കറാം	2	0	2	0	0.0002
കരി	653	7	229	50	0.070
കരിം	24	17	1	0	0.0026
കരു	15	1	4	4	0.0016
കരും	2	2	0	0	0.0002
കൈ	39	0	2	37	0.004
കേ	490	189	45	137	0.053
കേറ	45	4	17	1	0.0048
കൈ	104	0	31	0	0.011
കൈം	244	0	2	224	0.026
കൈ	95	0	30	6	0.010
കൈറ	12	1	2	0	0.0013
കൈം	68	0	0	68	0.0073

ആകെ പരിശോധിച്ച അക്ഷരങ്ങൾ

928971

അക്ഷരം	ആവൃത്തി	ആദ്യം	മധ്യം	അന്ത്യം	ശതമാനം
ക്ര	26	0	11	5	0.0028
ക്ര	5	0	3	1	0.0005
ക്രി	2	0	0	0	0.0002
ക്രീ	26	0	5	0	0.0028
ക്ര	5	0	2	3	0.0005

ഗ്ലം	18	0	0	18	0.0019
ഗ്ലി	54	0	45	6	0.0058
ഗ്ലാ	27	13	8	0	0.0029
ഗ്ലി	2	0	0	0	0.0002
ഗ്ലീ	156	0	141	1	0.017
ഗ്ലൂ	59	0	49	7	0.006

ആകെ പരിശോധിച്ച അക്ഷരങ്ങൾ

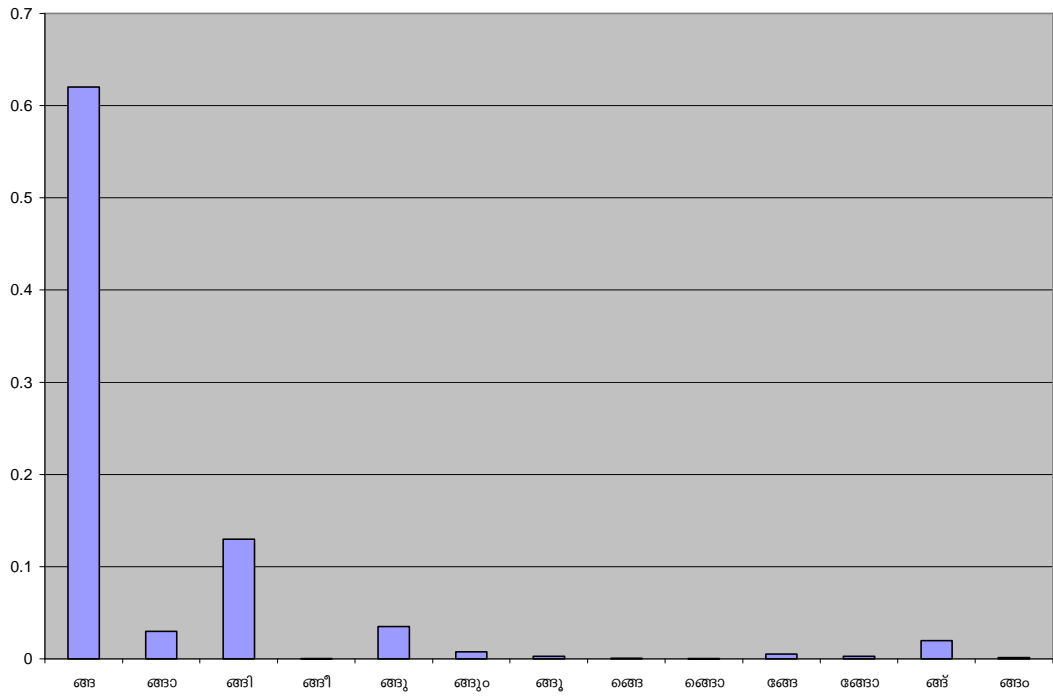
928971

അക്ഷരം	ആവൃത്തി	ആദ്യം	മധ്യം	അന്ത്യം	ശതമാനം
ക	720	0	282	24	0.078
കാ	110	0	42	1	0.012
കാം	1	0	1	0	0.0001
കി	941	1	105	4	0.101
കിം	6	0	6	0	0.0006
കീ	9	0	3	0	0.00097
കു	120	1	42	3	0.013
കും	11	0	0	11	0.0012
കു	1	0	1	0	0.00011
കെ	207	0	83	4	0.022
കൊ	12	0	4	2	0.0013
കേ	106	0	55	0	0.011
കോ	58	0	28	13	0.0062
ക്	196	0	4	188	0.021
കം	15	0	0	14	0.0016

ആകെ പരിശോധിച്ച അക്ഷരങ്ങൾ

928971

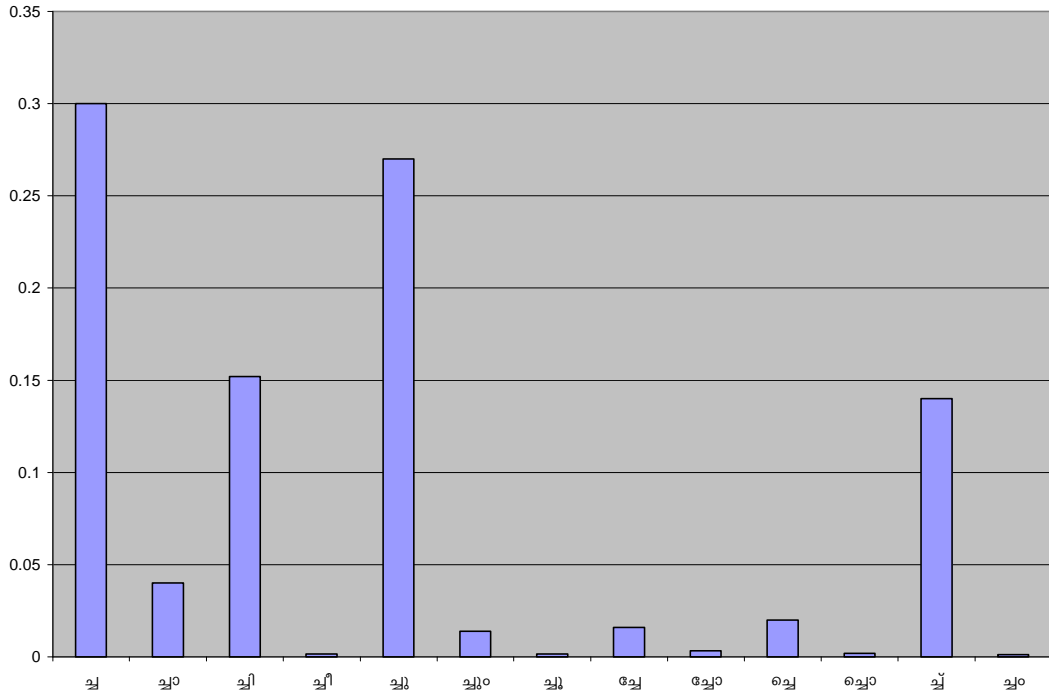
അക്ഷരം	ആവൃത്തി	ആദ്യം	മധ്യം	അന്ത്യം	ശതമാനം
ങ്ങ	5748	3	1443	15	0.62
ങ്ങാ	277	1	85	2	0.030
ങ്ങാം	12	0	0	12	0.0013
ങ്ങി	1173	2	318	243	0.13
ങ്ങിം	2	0	0	0	0.0002
ങ്ങീ	3	0	2	1	0.0003
ങ്ങു	327	0	145	12	0.035
ങ്ങും	71	0	1	70	0.0076
ങ്ങു	27	0	4	14	0.0029
ങ്ങെ	8	0	3	1	0.0008
ങ്ങൊ	2	0	2	0	0.0002
ങ്ങേ	46	0	31	6	0.0050
ങ്ങോ	25	0	9	4	0.0027
ങ്ങ്	183	1	2	175	0.020
ങ്ങം	13	0	0	13	0.0014
ങ്ങ്യം	1	0	0	0	0.0001



ആകെ പരിശോധിച്ച അക്ഷരങ്ങൾ

928971

അക്ഷരം	ആവൃത്തി	ആദ്യം	മധ്യം	അന്ത്യം	ശതമാനം
ച്ച	2749	1	614	1033	0.30
ച്ചറ	369	0	71	20	0.040
ച്ചി	1415	2	504	194	0.152
ച്ചിം	3	0	3	0	0.0003
ച്ചീ	15	0	7	0	0.0016
ച്ചീം	1	0	0	1	0.0001
ച്ചു	2554	2	239	1987	0.27
ച്ചും	132	0	3	127	0.014
ച്ചു	14	0	9	2	0.0015
ച്ചേ	146	0	59	5	0.016
ച്ചേറ	31	0	7	9	0.0033
ച്ചെ	182	0	56	18	0.02
ച്ചെറ	18	0	5	0	0.0019
ച്ച	1328	0	16	1303	0.14
ച്ചം	11	0	1	6	0.0012
ച്ച	1	0	0	0	0.0001



ആകെ പരിശോധിച്ച അക്ഷരങ്ങൾ

928971

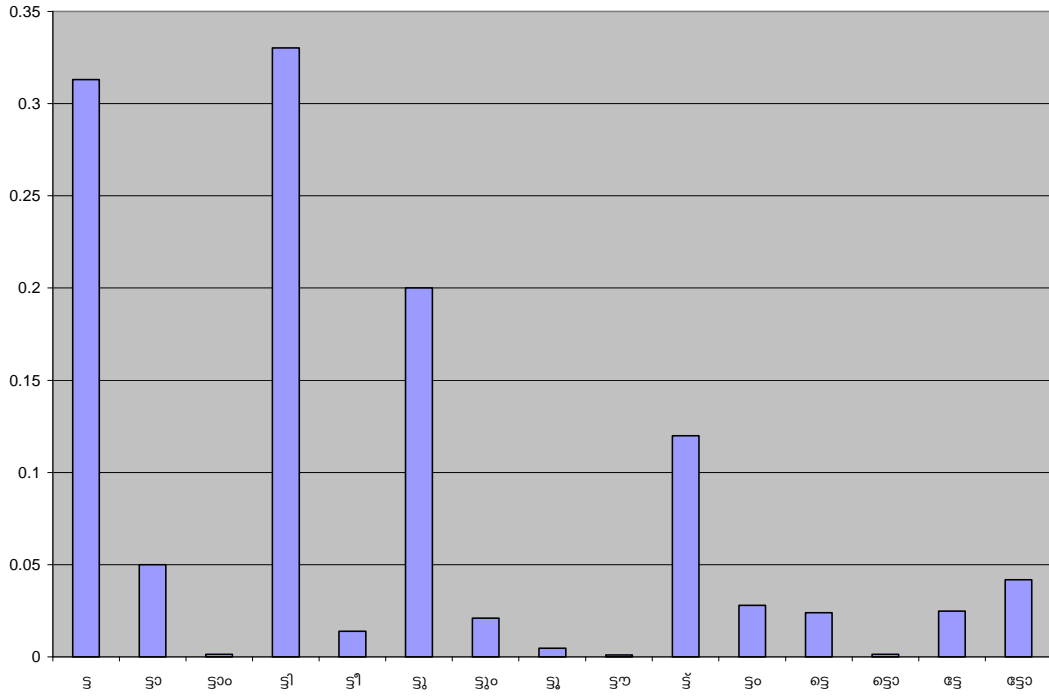
അക്ഷരം	ആവൃത്തി	ആദ്യം	മധ്യം	അന്ത്യം	ശതമാനം
ഞ	828	0	175	456	0.089
ഞാ	206	0	41	0	0.022
ഞാഠ	1	0	1	0	0.0001
ഞതി	335	3	163	4	0.036
ഞതു	1178	2	94	981	0.127
ഞതുഠ	4	0	0	4	0.0004
ഞതു	10	0	5	0	0.0011
ഞ്ഞ	293	1	184	0	0.032
ഞ്ഞാ	7	0	5	0	0.0007
ഞ്ഞേ	8	0	5	1	0.0008
ഞ്ഞോ	8	0	0	3	0.0008
ഞ്ഞി	167	0	3	161	0.018
ഞ്ഞഠ	37	0	2	35	0.0040

ആകെ പരിശോധിച്ച അക്ഷരങ്ങൾ

928971



അക്ഷരം	ആവൃത്തി	ആദ്യം	മധ്യം	അന്ത്യം	ശതമാനം
ട	2910	13	1024	547	0.313
ട്ര	465	18	164	7	0.050
ട്രം	12	0	2	8	0.0013
ട്രി	3064	1	1059	584	0.33
ട്രിം	8	0	2	1	0.0008
ട്രീ	129	0	71	6	0.014
ട്രൂ	1867	0	377	531	0.20
ട്രൂം	195	0	2	192	0.021
ട്രൂ	43	4	14	9	0.0046
ട്രൂം	1	0	0	1	0.0001
ട്രൂ	10	3	4	2	0.001
ട്രൂ	1106	0	53	990	0.12
ട്രൂം	259	0	0	256	0.028
ട്രൂ	223	8	43	133	0.024
ട്രൂം	12	0	4	4	0.0013
ട്രൂ	234	0	121	9	0.025
ട്രൂം	392	37	105	133	0.042



ആകെ പരിശോധിച്ച അക്ഷരങ്ങൾ

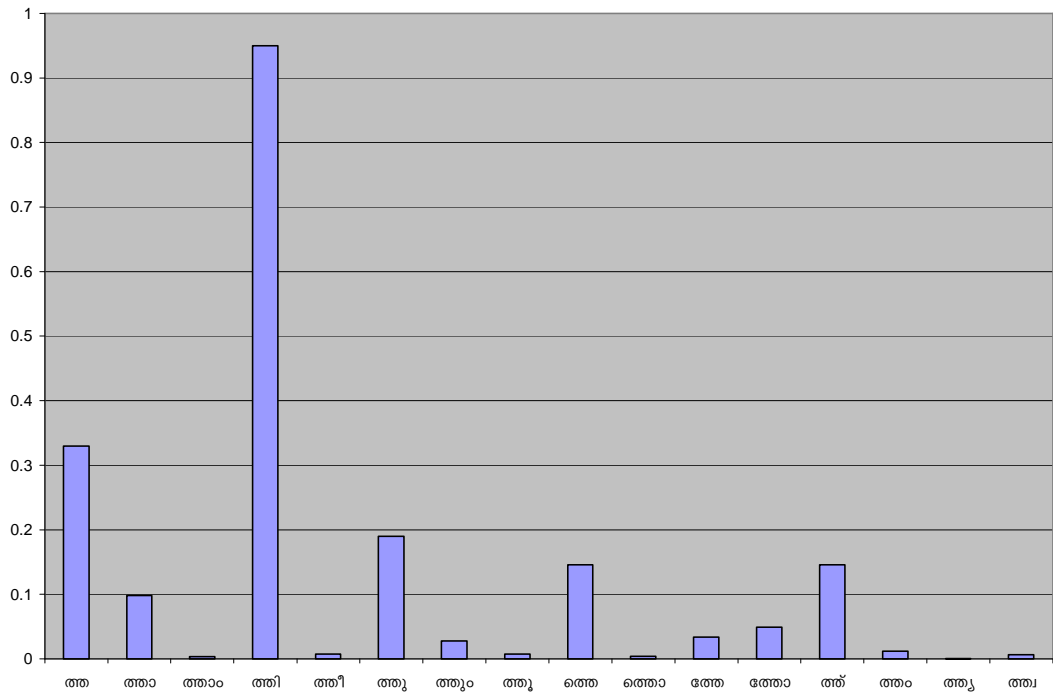
928971

അക്ഷരം	ആവൃത്തി	ആദ്യം	മധ്യം	അന്ത്യം	ശതമാനം
ണ്ട	1024	0	307	258	0.11
ണ്ടാ	1360	1	384	17	0.15
ണ്ടാം	166	0	0	152	0.018
ണ്ടി	850	0	301	222	0.091
ണ്ടിം	12	0	0	0	0.0013
ണ്ടീ	35	0	9	2	0.0038
ണ്ടു	822	0	139	416	0.088
ണ്ടും	238	0	0	238	0.026
ണ്ടു	6	0	0	4	0.0006
ണ്ടെ	422	0	131	7	0.045
ണ്ടോ	10	0	4	0	0.0010
ണ്ടേ	34	0	15	16	0.0037
ണ്ടോ	113	1	10	70	0.012
ണ്ടി	2261	1	28	2213	0.24
ണ്ടം	15	0	2	13	0.0016
ണ്ടു	2	0	0	0	0.0002
ണ്ടു	1	0	1	0	0.0001

ആകെ പരിശോധിച്ച അക്ഷരങ്ങൾ

928971

അക്ഷരം	ആവൃത്തി	ആദ്യം	മധ്യം	അന്ത്യം	ശതമാനം
ത	3028	3	929	607	0.33
താ	914	0	261	12	0.098
താം	32	1	3	27	0.0034
തീ	8805	23	1288	701	0.95
തീ	68	1	22	0	0.0073
തു	1789	1	455	502	0.19
തും	262	0	2	260	0.028
തു	67	2	17	3	0.0072
തും	1	0	0	1	0.0001
തെ	1355	1	137	999	0.146
തൊ	35	0	16	0	0.0038
തേ	317	1	99	13	0.034
തൊ	456	12	79	17	0.049
തൊം	6	0	2	4	0.0006
തൂ	4	0	4	0	0.0004
തീ	1357	0	7	1327	0.146
തം	113	0	2	110	0.012
തൂ	7	0	1	1	0.0007
തൂാ	1	0	0	1	0.0001
തൂം	1	0	0	1	0.0001
താ	60	0	43	4	0.0065
താം	1	0	0	1	0.0001



ആകെ പരിശോധിച്ച അക്ഷരങ്ങൾ

928971

അക്ഷരം	ആവൃത്തി	ആദ്യം	മധ്യം	അന്ത്യം	ശതമാനം
ത	192	0	54	44	0.021
താ	28	0	8	0	0.0030
തമി	167	0	64	29	0.018
തമഠ	71	0	1	69	0.0076
തമ്യ	16	0	8	2	0.0017
തമ്യഠ	3	0	0	3	0.0003

അക്ഷരം	ആവൃത്തി	ആദ്യം	മധ്യം	അന്ത്യം	ശതമാനം
ദ	31	0	9	4	0.0033
ദാ	41	0	25	0	0.0044
ദാഠ	13	0	1	10	0.0014
ദി	17	0	12	0	0.0018
ദീ	78	0	9	0	0.008
ദു	1	0	0	0	0.0001
ദേ	669	0	447	1	0.072
ദേഠ	4	0	2	0	0.0004
ദേ്യഠ	5	0	0	0	0.0005
ദു	2	0	0	1	0.0002
ദഠ	15	0	0	15	0.0016
ദ്യ	3	0	0	0	0.0003

ആകെ പരിശോധിച്ച അക്ഷരങ്ങൾ

928971

അക്ഷരം	ആവൃത്തി	ആദ്യം	മധ്യം	അന്ത്യം	ശതമാനം
ല	756	1	308	74	0.081
ലാ	97	0	40	0	0.0104
ലി	192	0	83	20	0.021
ലീ	97	0	29	0	0.010
ലു	1	0	0	0	0.0001
ലേ	32	0	8	0	0.0034
ലേ്യ	1	0	0	1	0.0001
ലേ്യാ	4	0	1	0	0.0004
ലം	12	0	0	11	0.0013
ല്യ	106	0	48	2	0.011
ല്യാ	33	0	9	0	0.0036
ല്യം	14	0	0	14	0.0015
ല്യാ	4	0	3	0	0.0004

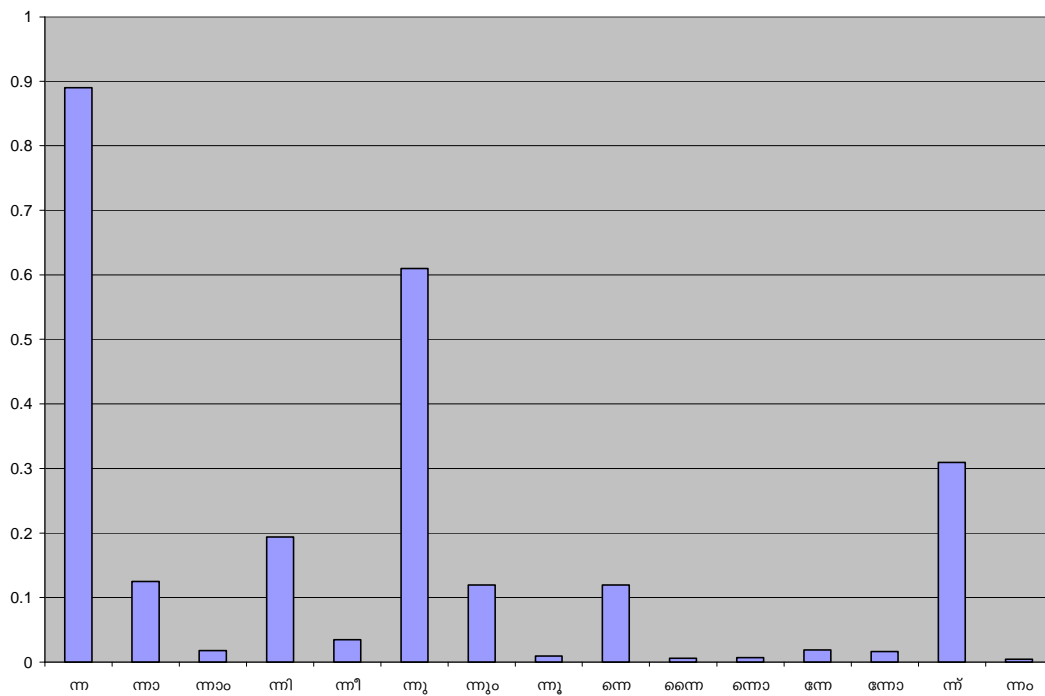
അക്ഷരം	ആവൃത്തി	ആദ്യം	മധ്യം	അന്ത്യം	ശതമാനം
ന	1898	1	438	79	0.204
നാ	197	0	73	36	0.021
നാം	1	0	0	0	0.0001
നി	321	0	101	86	0.035
നീ	186	1	55	46	0.020
നൂ	5	0	0	5	0.0005
നൂ	31	0	20	0	0.0033
നൈ	52	0	34	0	0.0056
നൈാ	21	0	11	0	0.002
നം	238	0	0	236	0.026
ന്യ	640	0	385	174	0.069
ന്യാ	57	0	1	47	0.0061
ന്യം	5	0	0	5	0.0005

അക്ഷരം	ആവൃത്തി	ആദ്യം	മധ്യം	അന്ത്യം	ശതമാനം
ന്ദ	585	0	135	32	0.063
ന്ദാ	22	0	6	1	0.0024
ന്ദി	180	0	73	34	0.019
ന്ദീ	4	0	4	0	0.0004
ന്ദൂ	126	0	31	70	0.014
ന്ദൂം	2	0	0	2	0.0002
ന്ദൂ	1	0	1	0	0.0001
ന്ദേ	123	0	49	0	0.013
ന്ദേ	48	0	2	46	0.0052
ന്ദം	15	0	0	15	0.0016
ന്ദ്യ	6	0	0	3	0.0006
ന്ദ്യം	6	0	0	6	0.0006

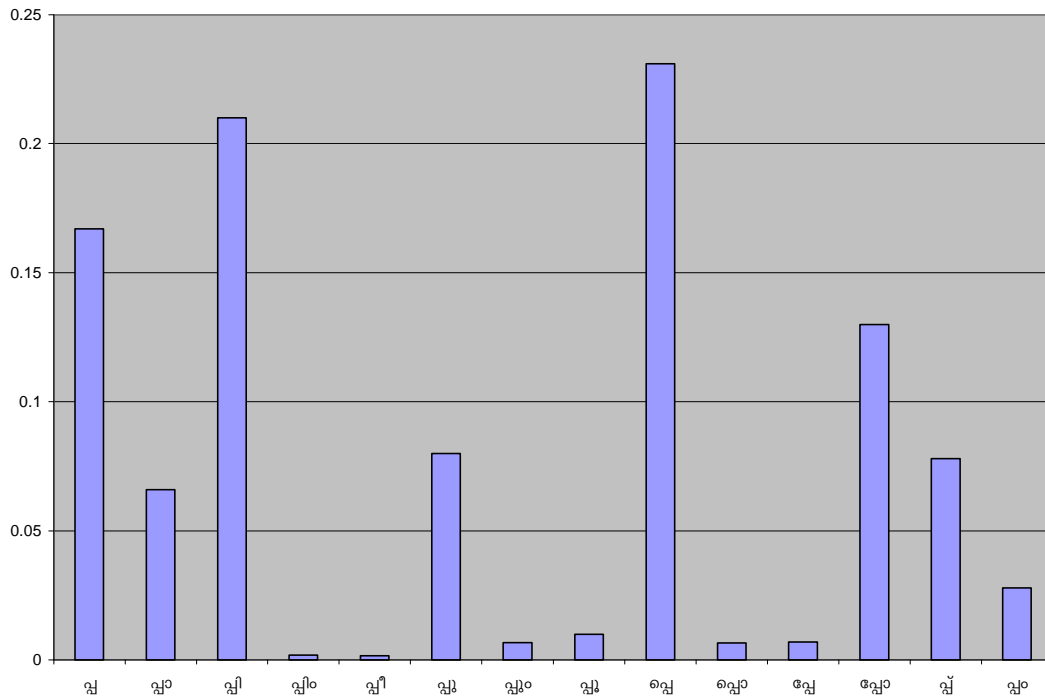
ആകെ പരിശോധിച്ച അക്ഷരങ്ങൾ

928971

അക്ഷരം	ആവൃത്തി	ആദ്യം	മധ്യം	അന്ത്യം	ശതമാനം
ന	8258	3	1146	3788	0.89
നാ	1162	10	583	26	0.125
നാം	168	0	4	161	0.0181
നീ	1803	1	818	84	0.194
നീ	327	0	153	154	0.035
നീം	1	0	0	1	0.00010
നു	5670	6	377	4562	0.610
നൂം	1154	1	3	1137	0.12
നു	90	1	15	2	0.0097
നൂം	1	0	0	1	0.0001
നെ	1037	1	173	656	0.12
നൈ	54	0	16	36	0.0058
നൊ	65	0	52	0	0.0070
നേ	178	0	56	31	0.019
നോ	153	1	58	47	0.0165
നോം	1	0	0	1	0.0001
ന്	2847	2	17	2814	0.31
നം	42	0	9	29	0.0045
ന്യ	9	0	7	0	0.00009
ന്യം	6	0	0	0	0.00006
ന്വ	1	0	0	0	0.0001



അക്ഷരം	ആവൃത്തി	ആദ്യം	മധ്യം	അന്ത്യം	ശതമാനം
പ	1552	3	641	31	0.167
പാ	609	1	239	4	0.066
പി	1956	11	693	45	0.21
പിം	18	0	18	0	0.0019
പിീ	15	0	6	0	0.0016
പു	738	2	207	127	0.080
പും	62	0	2	59	0.0067
പു	95	0	35	1	0.010
പുെ	2145	4	722	1	0.231
പുൊ	61	1	23	5	0.0066
പുേ	67	0	39	2	0.007
പുൊ	1166	5	552	16	0.13
പു	2	0	1	1	0.0002
പ	725	0	10	694	0.078
പം	259	1	1	256	0.028
പ	2	0	0	2	0.0002
പാ	5	0	1	0	0.00054
പ	12	0	12	0	0.0013



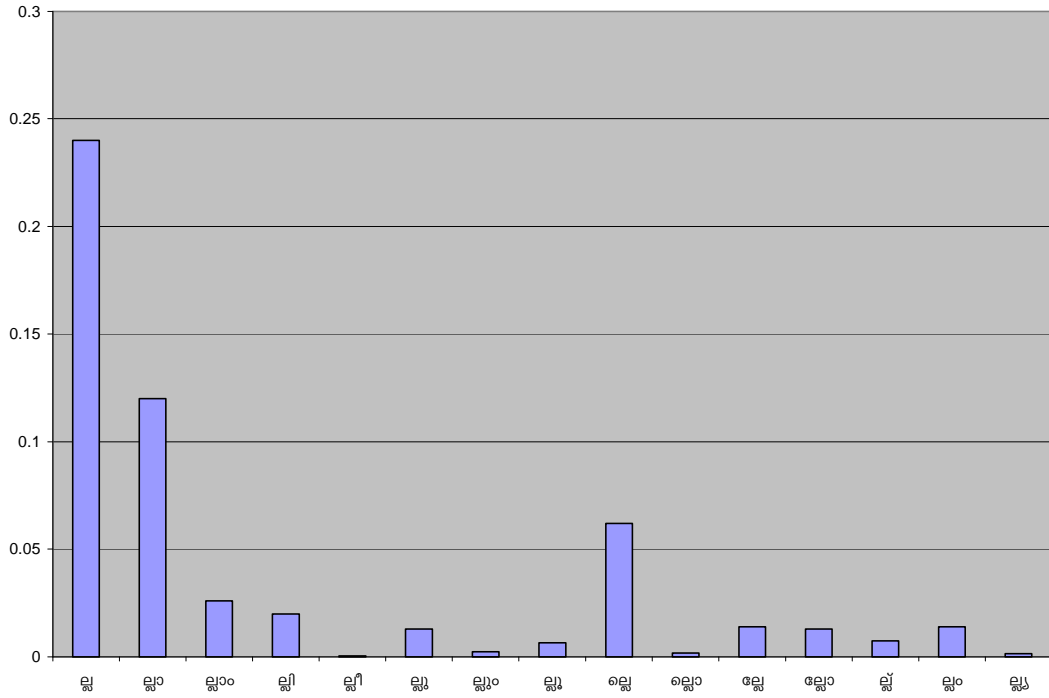
ആകെ പരിശോധിച്ച അക്ഷരങ്ങൾ

928971

അക്ഷരം	ആവൃത്തി	ആദ്യം	മധ്യം	അന്ത്യം	ശതമാനം
ശ്ര	296	0	180	22	0.032
ശ്രാ	426	0	135	8	0.046
ശ്രാം	8	0	0	8	0.0008
ശ്രീ	44	0	33	1	0.0047
ശ്രു	448	0	295	0	0.048
ശ്രും	111	0	0	110	0.012
ശ്രു	4	0	2	2	0.0004
ശ്ര്യ	11	0	1	7	0.0012
ശ്ര്യം	3	0	0	3	0.0003
ശ്ര്യം	7	0	1	6	0.00075
ശ്ര്യ	3	0	3	0	0.0003
ശ്രയാ	2	0	0	0	0.0002
ശ്ര്യേ	57	0	27	0	0.006
ശ്ര്യോ	7	0	2	5	0.0007

അക്ഷരം	ആവൃത്തി	ആദ്യം	മധ്യം	അന്ത്യം	ശതമാനം
ശ്ല	2207	1	292	1652	0.24
ശ്ലാ	1126	1	252	425	0.12
ശ്ലാം	237	1	0	236	0.026
ശ്ലീ	186	0	97	44	0.020
ശ്ലീം	2	0	2	0	0.0002
ശ്ലീ	4	0	1	1	0.0004
ശ്ലു	124	0	56	13	0.013
ശ്ലും	21	0	2	17	0.0023
ശ്ലു	61	0	17	1	0.0066
ശ്ലു	572	0	117	7	0.062
ശ്ലയാ	16	0	14	0	0.0017
ശ്ലേ	133	1	52	69	0.014
ശ്ലോ	124	0	3	120	0.013
ശ്ലേ	6	0	4	0	0.0006
ശ്ലേ	70	0	0	67	0.0075
ശ്ലം	134	1	2	127	0.014
ശ്ല	14	0	14	0	0.0015
ശ്ലാ	2	0	2	0	0.0002
ശ്ലാം	3	0	0	3	0.0003





ആകെ പരിശോധിച്ച അക്ഷരങ്ങൾ

928971

അക്ഷരം	ആവൃത്തി	ആദ്യം	മധ്യം	അന്ത്യം	ശതമാനം
ഇ	2208	0	145	1576	0.24
ഇാ	27	0	16	1	0.0030
ഇാം	3	0	0	3	0.0003
ഇി	366	0	109	84	0.039
ഇു	104	0	29	27	0.011
ഇും	10	0	0	10	0.0011
ഇു	103	0	16	75	0.011
ഇെ	3	0	0	0	0.0003
ഇേ	7	0	0	1	0.0007
ഇോ	3	0	2	1	0.0003
ഇേ	1	0	1	0	0.0001
ഇ	3	0	0	3	0.0003
ഇം	96	0	10	81	0.010

അക്ഷരം	ആവൃത്തി	ആദ്യം	മധ്യം	അന്ത്യം	ശതമാനം
വ	126	0	44	24	0.014
വാ	28	0	23	3	0.0030
വി	1	0	1	0	0.0001
വീ	41	0	8	0	0.0044
വു	2	0	0	0	0.0002
വെ	7	0	5	2	0.0007
വേ	32	0	12	4	0.0034
വം	12	0	0	12	0.0013

ആകെ പരിശോധിച്ച അക്ഷരങ്ങൾ

928971

അക്ഷരം	ആവൃത്തി	ആദ്യം	മധ്യം	അന്ത്യം	ശതമാനം
സ്	373	8	182	3	0.040
സാ	242	194	13	2	0.026
സി	269	7	183	6	0.029
സിം	3	0	3	0	0.0003
സീ	33	27	4	0	0.0036
സീം	1	0	0	1	0.0001
സു	67	10	20	5	0.0072
സും	8	0	0	8	0.0008
സു	6	2	4	0	0.0006
സ്യ	31	26	5	0	0.0033
സ്യ	38	13	20	0	0.0041
സ്യ	174	106	50	0	0.019
സ്യോ	91	56	14	2	0.0098
സു	4	4	0	0	0.0004
സു	1	1	0	0	0.0001
സ്	525	1	6	497	0.057
സം	16	0	9	7	0.0017
സ്യ	14	0	8	0	0.0015

അക്ഷരം	ആവൃത്തി	ആദ്യം	മധ്യം	അന്ത്യം	ശതമാനം
ച	82	0	33	3	0.0088
ചാ	9	0	2	4	0.0009
ചി	3	0	3	0	0.0003
ചം	1	0	0	1	0.0001
ചാ	1	0	0	0	0.0001

അക്ഷരം	ആവൃത്തി	ആദ്യം	മധ്യം	അന്ത്യം	ശതമാനം
സ	514	18	137	14	0.055
സാ	5	0	2	0	0.0005
സി	382	2	47	61	0.041
സു	46	0	14	17	0.0050
സു	3	0	0	3	0.0003
സം	43	5	0	36	0.0046
സ്യ	32	0	9	16	0.0034
സ്യ	9	0	2	0	0.0009

അക്ഷരം	ആവൃത്തി	ആദ്യം	മധ്യം	അന്ത്യം	ശതമാനം
സ	354	0	145	38	0.038
സാ	45	0	20	5	0.0048
സി	14	0	5	0	0.0015
സു	4	0	4	0	0.00004
സം	2	0	0	2	0.0002
സ്യ	23	0	11	7	0.0025
സ്യ	2	0	0	0	0.0002
സ്യം	1	0	0	1	0.0001
സ്യ	1	0	0	0	0.0001

ആകെ പരിശോധിച്ച അക്ഷരങ്ങൾ

928971

അക്ഷരം	ആവൃത്തി	ആദ്യം	മധ്യം	അന്ത്യം	ശതമാനം
ജ	29	0	13	0	0.003
ജാ	2	0	1	1	0.0002
ജി	17	0	6	0	0.0018
ജീ	18	0	0	0	0.0020
ജു	1	0	0	0	0.0001
ജും	6	0	0	6	0.0006
ജ്ജ	50	0	3	44	0.0054
ജ്ജ	2	0	0	0	0.0002

അക്ഷരം	ആവൃത്തി	ആദ്യം	മധ്യം	അന്ത്യം	ശതമാനം
ഞ	55	0	17	10	0.0059
ഞാ	76	15	21	0	0.008
ഞി	6	0	1	2	0.0006
ഞു	9	0	2	4	0.0006

അക്ഷരം	ആവൃത്തി	ആദ്യം	മധ്യം	അന്ത്യം	ശതമാനം
സ്ഥ	603	173	171	87	0.065
സ്ഥാ	1337	432	573	5	0.144
സ്ഥി	152	94	15	0	0.016
സ്ഥു	1	1	0	0	0.0001
സ്ഥ്	2	0	2	0	0.0002
സ്ഥ്യ	2	0	0	2	0.0002
സ്ഥ്യം	5	0	0	5	0.0005
ന്ഥ	40	0	21	0	0.0043
ന്ഥം	32	0	0	32	0.0034
ന്ഥ	2	0	0	1	0.0002

അക്ഷരം	ആവൃത്തി	ആദ്യം	മധ്യം	അന്ത്യം	ശതമാനം
ട	10	0	4	0	0.0011
ടാ	23	0	7	0	0.0025
ടി	2	0	0	0	0.0002

അക്ഷരം	ആവൃത്തി	ആദ്യം	മധ്യം	അന്ത്യം	ശതമാനം
ശ്വ	48	0	13	0	0.0052
ശ്വാ	32	0	20	0	0.0034
ശ്വി	44	0	23	0	0.0047

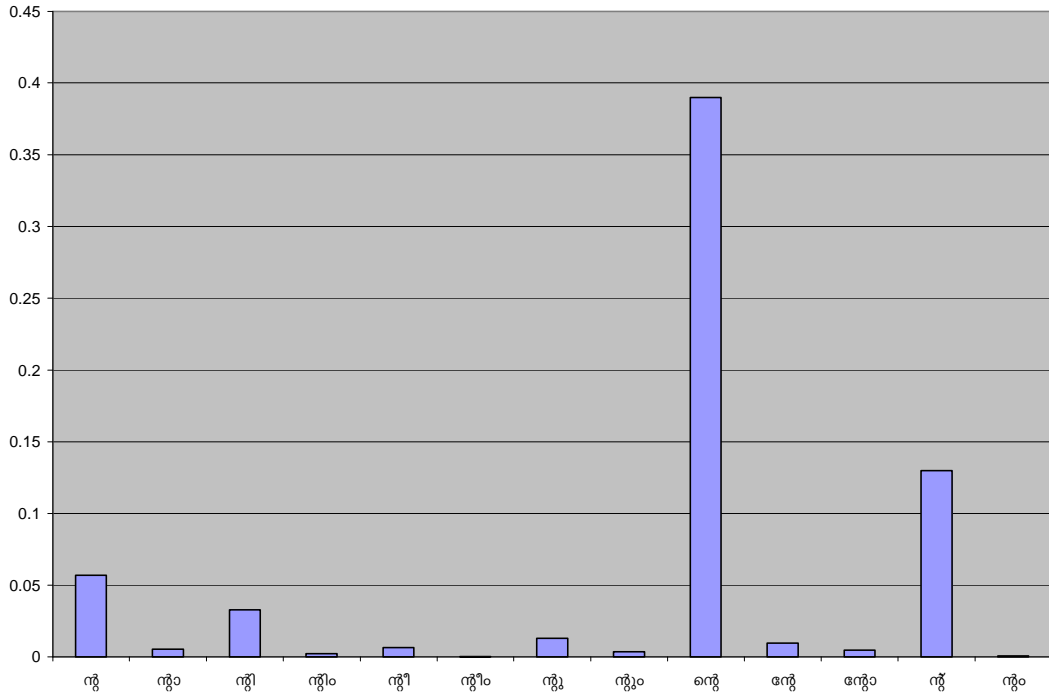
അക്ഷരം	ആവൃത്തി	ആദ്യം	മധ്യം	അന്ത്യം	ശതമാനം
ശ്മ	165	0	43	17	0.018
ശ്മാ	52	0	33	4	0.0056
ശ്മി	21	0	7	0	0.002
ശ്മീ	2	0	2	0	0.0002
ശ്മ്യ	4	0	0	2	0.0004
ശ്മ്യം	7	0	0	7	0.0007

തോ	3	0	0	0	0.0003
----	---	---	---	---	--------

ആകെ പരിശോധിച്ച അക്ഷരങ്ങൾ 928971

അക്ഷരം	ആവൃത്തി	ആദ്യം	മധ്യം	അന്ത്യം	ശതമാനം
ക	441	0	225	29	0.047
കാ	63	0	35	0	0.0068
കി	350	0	119	102	0.038
കു	2	0	0	0	0.0002
കൂ	6	0	0	6	0.0006
കെ	3	0	3	0	0.0003
കേ	11	0	0	11	0.0012

അക്ഷരം	ആവൃത്തി	ആദ്യം	മധ്യം	അന്ത്യം	ശതമാനം
ന്	525	2	272	26	0.057
നാ	48	0	8	0	0.0052
നി	302	0	76	27	0.033
നീ	20	0	11	0	0.0022
നൂ	59	0	17	3	0.0064
നൂ	2	0	0	2	0.0002
നു	117	0	65	2	0.013
നൂ	33	0	0	33	0.0036
നെ	3587	11	9	3365	0.39
നേ	90	0	10	0	0.0097
നോ	43	0	4	7	0.0046
നൂ	1220	0	21	1135	0.13
നൂ	6	0	2	0	0.0006



ആകെ പരിശോധിച്ച അക്ഷരങ്ങൾ

928971

അക്ഷരം	ആവൃത്തി	ആദ്യം	മധ്യം	അന്ത്യം	ശതമാനം
റ	1305	3	640	92	0.14
റാ	196	3	76	31	0.021
റാം	1	0	0	1	0.0001
റി	1290	13	345	440	0.14
റിം	18	0	1	0	0.0019
റീ	55	5	2	1	0.0059
റു	558	0	187	163	0.060
റും	105	0	1	102	0.011
റു	24	0	18	1	0.0026
റു	2	0	0	2	0.0002
റു	1110	0	80	964	0.12
റു	135	0	51	78	0.015
റു	60	0	3	0	0.0064
റു	1	0	1	0	0.0001
റി	2	0	2	0	0.0002
ര	34	0	11	0	0.0037
രാ	105	0	84	0	0.012
ര	2	0	2	0	0.0002
ര	112	0	63	6	0.012
രാ	49	5	16	16	0.0053
രാ	1	0	0	1	0.0001

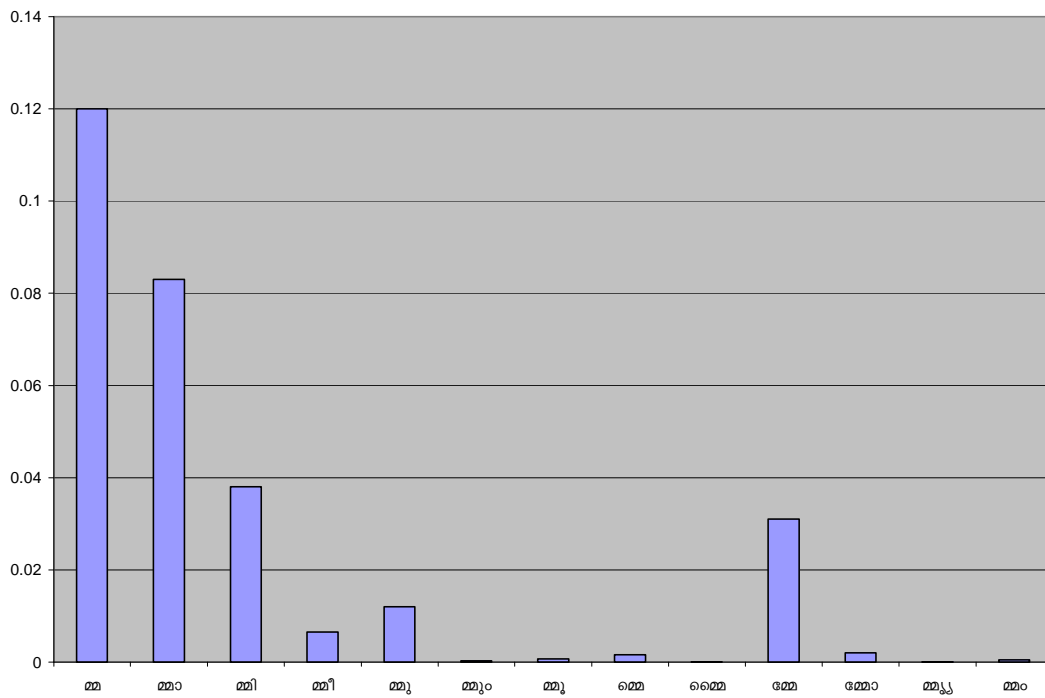
അക്ഷരം	ആവൃത്തി	ആദ്യം	മധ്യം	അന്ത്യം	ശതമാനം
ഞ്ച	238	5	114	2	0.026
ഞ്ചാ	304	0	183	0	0.033
ഞ്ചാം	23	0	3	16	0.0025
ഞ്ചി	418	0	81	11	0.045
ഞ്ചി്	1	0	1	0	0.0001
ഞ്ചീ	1	0	1	0	0.0001
ഞ്ചു	188	0	36	82	0.020
ഞ്ചൂ	12	4	0	0	0.0013
ഞ്ചെ	2	0	0	0	0.0002
ഞ്ചൈ	4	0	2	0	0.0004
ഞ്ചേ	72	0	30	0	0.0078
ഞ്ചോ	9	0	1	3	0.0009
ഞ്ചി്	177	0	2	170	0.019
ഞ്ചം	3	0	0	2	0.0003
ഞ്ച	7	0	4	0	0.0007

അക്ഷരം	ആവൃത്തി	ആദ്യം	മധ്യം	അന്ത്യം	ശതമാനം
മ്പ	1282	2	664	19	0.138
മ്പാ	159	0	107	0	0.017
മ്പാം	5	0	3	2	0.0005
മ്പി	248	0	116	25	0.027
മ്പിം	2	0	2	0	0.0002
മ്പീ	1	0	0	0	10.0001
മ്പൂ	113	1	41	26	0.012
മ്പൂം	13	0	0	13	0.0014
മ്പു	118	0	74	0	0.013
മ്പെ	9	0	3	3	0.00097
മ്പൈ	14	0	12	2	0.0015
മ്പൊ	13	0	5	3	0.0014
മ്പേ	11	0	6	2	0.0012
മ്പോ	438	0	24	2	0.047
മ്പൗ	3	0	2	0	0.0003
മ്പ്	180	0	4	174	0.019
മ്പം	7	0	2	5	0.0007
മ്പ്യ	127	0	46	0	0.014
മ്പ്യാ	26	0	18	0	0.0028
മ്പ്യൂ	3	0	2	0	0.0003
മ്പ്യൂ	113	0	100	0	0.012
മ്പ്ര	20	0	13	2	0.0022
മ്പ്രാ	2	0	0	0	0.0002
മ്പ്രം	2	0	0	2	0.0002

ആകെ പരിശോധിച്ച അക്ഷരങ്ങൾ

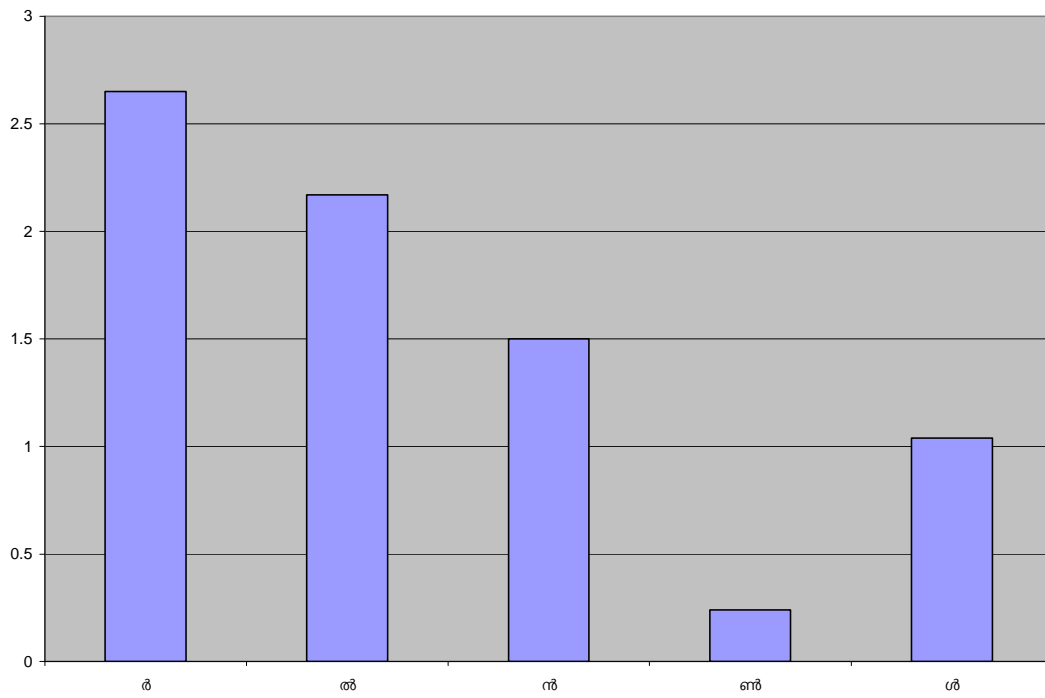
928971

അക്ഷരം	ആവൃത്തി	ആദ്യം	മധ്യം	അന്ത്യം	ശതമാനം
മ്മ	1088	0	264	549	0.12
മ്മാ	773	0	449	24	0.083
മ്മി	355	1	202	16	0.038
മ്മീ	60	0	26	0	0.0065
മ്മു	113	0	94	4	0.012
മ്മും	3	0	0	3	0.0003
മ്മു	7	0	4	0	0.0007
മ്മെ	15	0	3	12	0.0016
മ്മൈ	1	0	0	1	0.0001
മ്മേ	285	0	184	9	0.031
മ്മോ	18	0	2	3	0.0020
മ്മ്യ്യ	1	0	1	0	0.0001
മ്മം	5	0	0	5	0.0005



3.3.4 ചില്ലക്ഷരങ്ങൾ

അക്ഷരം	ആവൃത്തി	ആദ്യം	മധ്യം	അന്ത്യം	ശതമാനം
ർ	24622	0	6496	9245	2.65
ൽ	20206	0	1652	16239	2.17
ൻ	14372	0	1758	9865	1.5
ൺ	2195	0	561	1047	0.24
ൾ	9646	0	541	7168	1.038





### 3.4 വാക്കുകളുടെ ആവൃത്തി

ഡേറ്റയിൽ 50-ൽപരം തവണ ആവർത്തിച്ചുവന്നിട്ടുള്ള വാക്കുകൾ മാത്രമേ പട്ടികയിൽ ചേർത്തിട്ടുള്ളൂ. ഇംഗ്ലീഷ് പദങ്ങൾ വളരെയധികം വന്നിരിക്കുന്നത് ശ്രദ്ധിക്കേണ്ടിയിരിക്കുന്നു. അന്യഭാഷാ സ്വാധീനത്തെപ്പറ്റിയുള്ള വിശകലനം ഈ പട്ടികയുടെ അവസാനം കൊടുത്തിരിക്കുന്നത് ശ്രദ്ധിക്കുക.

ആകെ പരിശോധിച്ച വാക്കുകൾ 231031

അക്ഷരം	ആവൃത്തി
ഒരു	1162
ഈ	913
തിരുവനന്തപുരം	739
പറഞ്ഞു	694
എന്ന	489
രൂപ	414
ഫോൺ	386
സർക്കാർ	378
സമ്മാനം	375
ഭാര്യ	374
അദ്ദേഹം	368
തന്നെ	366
സംസ്ഥാന	365
കേരള	339
മക്കൾ	321
സെക്രട്ടറി	319
നിന്ന്	319
നായർ	293
വരെ	292
ചെയ്തു	282
മുതൽ	279
ഡോ	278
പുതിയ	278
ഉദ്ഘാടനം	267
മന്ത്രി	267
എന്നാൽ	262
പ്രസിഡന്റ്	258
ആ	249
മാത്രം	248
ഇന്ന്	240
കോട്ടയം	235
നിന്നും	232
കൂടുതൽ	227
ഇപ്പോൾ	217

നിര്യാതനായി	214
മരുമക്കൾ	208
രണ്ടു	208
ജനറൽ	206
കഴിഞ്ഞ	203
മാത്യുഭൂമി	201
രാവിലെ	199
അവർ	199
ജില്ലാ	197
വർഷം	196
അത്	193
വർഷ	192
പരേതനായ	191
സംസ്കാരം	189
എന്നിവർ	189
വില	186
എല്ലാ	185
ആൻഡ്	181
ഇന്ത്യൻ	180
ഓഫ്	177
എന്ന്	177
സ്വന്തം	175
അറിയിച്ചു	174
വേണം	173
ഏറ്റവും	173
ആഗസ്ത്	172
റോഡ്	169
നിന്നു	166
കോഴിക്കോട്	165
പറയുന്നു	164
രണ്ട്	164
പരിചയം	163
കൊണ്ട്	163
ജോലി	159
സമ്മേളനം	159
ചില	156
ഇത്	155
ഇന്ത്യ	155

വീണ്ടും	155
നടത്തി	154
വീട്ടിൽ	154
നൽകി	152
എന്നീ	152
എറണാകുളം	150
പദ്ധതി	149
യോഗ്യത	149
രണ്ടാം	149
നമ്പർ	149
ബന്ധപ്പെടുക	146
സമരം	146
കൊച്ചി	145
ദിവസം	140
മികച്ച	140
തുടങ്ങിയ	139
ഞാൻ	139
വിവിധ	139
കോളേജ്	138
പക്ഷേ	136
മുൻ	136
ഇവിടെ	134
തോമസ്	133
മൂന്നു	131
നടന്ന	131
യോഗം	130
തന്റെ	130
കൂടി	130
നല്ല	128
അപേക്ഷ	128
പോലീസ്	127
ഇംഗ്ലീഷ്	126
കവി	124
മെഡിക്കൽ	123
മരിച്ചു	121
കോൺഗ്രസ്	120
എന്നു	119

നിര്യാതയായി	118
സമീപം	118
ആവശ്യമുണ്ട്	117
പ്രത്യേക	117
ഇനി	117
അവരുടെ	116
ആദ്യ	116
ശേഷം	115
പല	114
മനോരമ	114
മീറ്റർ	114
റിട്ട	113
പ്രസംഗിച്ചു	112
ദേശീയ	112
ഓരോ	112
നേടി	112
ലക്ഷം	112
സർവീസ്	111
മലയാള	109
അസോസിയേഷൻ	108
തീയതി	108
ഒന്നാം	107
എന്നിവ	107
നടത്തിയ	107
പാർട്ടി	106
ഉടൻ	106
ആവശ്യപ്പെട്ടു	106
ഡയറക്ടർ	105
തിങ്കളാഴ്ച	105
അവസാന	105
കോഴ്സ്	105
പിന്നീട്	104
അതിന്റെ	104
നടക്കുന്ന	102
സെന്റർ	101
ഗവ	100

വലിയ	100
കൊല്ലം	100
രവി	100
നിങ്ങളുടെ	99
മാസം	99
അന്വേഷണം	99
റിപ്പോർട്ട്	98
രാത്രി	98
വൈസ്	98
അഡ്വ	98
പ്രവേശനം	98
ശ്രീ	97
കേരളത്തിൽ	97
യുടെ	96
വർഷത്തെ	96
ആന്റ്	96
പ്രസിഡണ്ട്	96
സഹിതം	96
പേർ	95
ബാങ്ക്	93
സയൻസ്	93
മൂന്ന്	93
പാലക്കാട്	93
കമ്പ്യൂട്ടർ	93
കുമാർ	92
പറയുന്നത്	91
ലഭിക്കും	91
സൂപ്പർ	91
നടത്തുന്ന	91
കേന്ദ്ര	91
ഫീസ്	91
കുറഞ്ഞ	91
വഴി	91
സൗജന്യം	90
പെൻഷൻ	89
അതു	89
ആണ്	88
മുഖ്യമന്ത്രി	88
വഹിച്ചു	87

ചെയ്ത	86
ലേഖകൻ	86
തൃശ്ശൂർ	86
എന്റെ	86
വിദ്യാഭ്യാസ	85
ഞായറാഴ്ച	85
മകൻ	85
മറ്റു	85
മാനേജർ	84
ബോർഡ്	84
കോടതി	84
സ്കൂൾ	83
നമ്മുടെ	82
ചെയർമാൻ	82
എഞ്ചിനീയറിംഗ്	81
മൊബൈൽ	81
ഇതു	81
പ്രമുഖ	80
രൂപയുടെ	80
വേണ്ടി	80
ഹോട്ടൽ	80
വിഭാഗത്തിൽ	80
മാത്രമേ	79
ബാബു	79
പോലെ	79
ആവശ്യം	79
സമയം	79
തുടങ്ങി	79
നറുക്കെടുപ്പ്	79
സീനിയർ	78
ശനിയാഴ്ച	78
മുമ്പ്	78
സർക്കാരിന്റെ	78
എന്നാണ്	78
പ്രധാന	78
കാരണം	77
എന്നും	77
ചെയ്യും	77
മുഴുവൻ	76
അദ്ദേഹത്തിന്റെ	76

ഇവ	75
നടന്നു	75
ഉയർന്ന	75
പഞ്ചായത്ത്	75
സ്പോർട്സ്	74
ഓഫീസ്	74
കമ്പനി	74
ഇതിൽ	74
അതിൽ	73
നടപടി	72
അടുത്ത	72
നൽകിയ	72
സൗകര്യം	72
അധ്യക്ഷത	72
ഏറെ	71
സമിതി	71
കേരളത്തിലെ	71
പേരിൽ	70
നൽകുന്ന	70
സുരേഷ്	70
പരിശീലനം	70
ആഗസ്റ്റ്	69
അഞ്ചു	69
വനിതാ	69
മറ്റ്	69
സഹകരണ	68
മുംബൈ	68
ചർച്ച	68
പഴയ	68
അമ്മ	67
ന്യൂഡൽഹി	67
കാര്യങ്ങൾ	67
വിദ്യാർഥി	67
തുടർന്ന്	67
സിനിമ	67
വിലാസത്തിൽ	67
കൗൺസിൽ	66
പ്രവർത്തനം	66
നടക്കും	66
പോലും	66
ആന്റണി	66

മാത്രമാണ്	66
സ്വകാര്യ	65
സ്വാശ്രയ	65
തൊഴിൽ	65
സാമ്പത്തിക	65
ഡിപ്ലോമ	65
വിതരണം	65
രാജൻ	64
ചെയ്യുന്നു	64
ഓണം	64
റാജ്	64
പരിചയവും	64
തൃശൂർ	64
വളരെ	64
കൂടാതെ	64
വന്നു	63
നേതൃത്വം	63
മലയാളം	63
വ്യക്തമാക്കി	63
നേടിയ	63
അവാർഡ്	62
അവിടെ	62
പരീക്ഷ	62
ജോസഫ്	62
ഇന്നലെ	62
തുക	62
ചെയ്യുന്ന	62
നാം	62
നേതാവ്	62
വൈദ്യുതി	61
ഇലക്ട്രോണിക്സ്	61
യൂണിയൻ	61
സമ്മാനങ്ങൾ	61
വരുന്ന	61
വ്യവസായ	61
നടത്തും	61



കോവളം	61
സൗജന്യ	60
മറ്റൊരു	60
ലഭ്യമാണ്	60
എഞ്ചിനീയർ	60
ഓഗസ്റ്റ്	60
ജോർജ്ജ്	60
സംഘം	60
ഇന്നു	60
നെടുമങ്ങാട്	60
സെന്റർ	60
ജേക്കബ്	59
അക്കാദമി	59
യോഗത്തിൽ	59
തുടരുന്നു	59
മാനേജ്മെന്റ്	59
വിജയൻ	59
സ്ഥലം	58
ത്യൂർ	58
ഓഫർ	58
ആദ്യം	58
ചെയ്തത്	58
പറയുന്ന	58
കവിത	58
പണം	58
ദ്	57
അസിസ്റ്റന്റ്	57
നിലയിൽ	57
നിയമനം	57
പ്രതിപക്ഷ	57
കോടി	57
ഇത്തരം	57
പിള്ള	57
ഭാഗമായി	56

അഡ്മിഷൻ	56
സംബന്ധിച്ച	56
അപേക്ഷിക്കാം	56
ചടങ്ങിൽ	56
വിദേശ	55
മാത്രമല്ല	55
സംവിധാനം	55
വൈകിട്ട്	55
ഓഫീസർ	55
ആറ്റിങ്ങൽ	55
പോയിന്റ്	55
വെള്ളം	55
ഫലം	55
താഴെ	54
ഭൂമി	54
നമ്പൂതിരി	54
മണിക്ക്	54
നാഷണൽ	54
അവൾ	54
ഞങ്ങളുടെ	54
നിയമം	54
നേരിൽ	54
നേരത്തെ	54
ഫെഡറേഷൻ	53
രംഗത്ത്	53
വേൾഡ്	53
സഞ്ചയനം	53
നേതൃത്വത്തിൽ	53
വിലാസം	53
സ്ഥാനം	53
ചൂണ്ടിക്കാട്ടി	53
തന്നെയാണ്	53

കഴിഞ്ഞു	53
തിരഞ്ഞെടുത്തു	53
ശതമാനം	52
ആലപ്പുഴ	52
വിഭാഗം	52
ടെൻഡർ	52
പോയി	52
തുടങ്ങിയവർ	52
ഫുട്ബോൾ	52
താൻ	52
മകൾ	52
ദർഘാസ്	51
ഇന്ത്യയുടെ	51
പ്രകാരം	51
ചെയ്യുന്നത്	51
അധികൃതർ	51
മറുപടി	51
സിവിൽ	51
ഇലക്ട്രിക്കൽ	51
പിള്ള	51
അഞ്ച്	51
എയർ	51
ഫോട്ടോ	51
ശമ്പളം	51
കോൺഗ്രസ്സ്	51
അവിശ്വാസ	51

ദാതുഭൂമി, മലയാള മനോരമ എന്നീ പത്രങ്ങളുടെ തിരുവനന്തപുരം എഡിഷൻകളിൽ നിന്നുള്ള പേജുകൾ വിശകലനത്തായി ഉപയോഗിച്ചതുമൂലം സ്ഥലനാമങ്ങൾ, പേരുകൾ എന്നിവ കൂടുതലായി ആവർത്തിച്ചുവന്നിട്ടുണ്ട്.

### 3.4.1 അന്യഭാഷകളുടെ സ്വാധീനം

ആശയവിനിമയോപാധിയായ ഭാഷയിൽ, വിഭിന്ന ഭാഷണ സമൂഹങ്ങളിൽപ്പെട്ട വ്യക്തികളുടെ അന്യോന്യസമ്പർക്കത്തിന്റെ ഫലമായി ഭാഷാസങ്കലനം ഉണ്ടാകാറുണ്ട്. ഓരോ ഭാഷണസമൂഹവും തങ്ങളുടെ അയൽ ഭാഷണസമൂഹങ്ങളിൽ നിന്ന് പലതും സ്വീകരിക്കുന്നതിന്റെ ഭാഗമായി ഒരു ഭാഷയുടെ സ്വനിമം, രൂപിമം, പദസമ്പത്ത് തുടങ്ങിയവ മറ്റ് ഭാഷയിലേക്ക് സ്വീകരിക്കപ്പെടുന്നു. ഇതിൽ ഏറ്റവും കൂടുതൽ സങ്കലനം നടക്കുന്നത് പദതലത്തിലായിരിക്കും. മറ്റ് ഭാഷകളിൽ നിന്ന് സ്വീകരിക്കപ്പെടുന്ന പദങ്ങളാണ് പരകീയ പദങ്ങൾ.

ഒരു ദിഭാഷകന് അന്യഭാഷാപദങ്ങളെ പ്രാഥമികഭാഷയിലേക്ക് ആദാനം ചെയ്യാൻ എളുപ്പം സാധിക്കും. അാന്യതാബോധം കൊണ്ടും അവശ്യബോധം കൊണ്ടും അയാൾ മേൻമയേറിയ അംശങ്ങളെ ഉൾക്കൊള്ളുന്നു.

50-ൽ കൂടുതൽ ആവൃത്തിയുള്ള ആകെ 394 വാക്കുകളിൽ 66 എണ്ണം വിദേശ ഭാഷകളിൽ നിന്ന് കടം കൊണ്ടിട്ടുള്ളവയാണ്. ചിലപദങ്ങളുടെ ഉപയോഗക്കൂടുതൽ കൊണ്ട് അവ അന്യഭാഷാപദങ്ങളാണെന്ന് തിരിച്ചറിയാൻ തന്നെ പ്രയാസമുണ്ട്.. ഏക ഭാഷികനായ ഒരാളിനെ സംബന്ധിച്ചിടത്തോളം താൻ ഉപയോഗിക്കുന്ന എല്ലാപദങ്ങളും അയാൾക്ക് മാതൃഭാഷയിൽ നിന്ന് ലഭിച്ചതായിരിക്കും. അന്യഭാഷാപദങ്ങൾ സ്വഭാഷാലിപിയിലാക്കി ഉപയോഗിക്കുന്നതിനാൽ വിദ്യാഭ്യാസം കുറഞ്ഞവർക്കുകൂടി ഇവ സുപരിചിതമാണ്. അതിനിടയിൽ പരകീയപദങ്ങൾ ഉണ്ടെങ്കിൽ തന്നെ അവയുടെ ആഗതസ്വഭാവം അയാൾക്ക് ശ്രദ്ധേയമായിത്തീരുന്നില്ല. ഭാഷ കൈകാര്യം ചെയ്യുന്ന ഭൂരിപക്ഷം ആളുകൾക്കും തങ്ങൾ ഉപയോഗിക്കുന്ന ഭാഷാമൂലകങ്ങളിൽ സ്വകീയം ഏതൊക്കെയാണ് പരകീയം ഏതൊക്കെയാണ് എന്ന് നിശ്ചയമുണ്ടാവില്ല. ഇവിടെ പട്ടികയിൽ ചേർത്തിരിക്കുന്ന വാക്കുകൾ പരിശോധിച്ചാൽ മലയാളത്തിന്റെ സ്വന്തമെന്ന് പറയാൻ സാധിക്കുന്നത് വളരെ കുറച്ച് മാത്രം. കടംകുട്ടിയായ സംസ്കൃതപദങ്ങൾ അധികം കടംകൊണ്ട് അർത്ഥം മനസ്സിലാക്കാൻ ഡിക്ഷണറി തപ്പി നടക്കുന്നതിനേക്കാൾ എന്തുകൊണ്ടും യോജിച്ചതാണ് ഈ പദങ്ങൾ.

വാക്ക്	ആവൃത്തി
ഫോൺ	386
സെക്രട്ടറി	319
ഡോ	278
പ്രസിഡന്റ്	258
ജനറൽ	206
ആൻഡ്	181
ഇന്ത്യൻ	180
ഓഫ്	177
റോഡ്	169
നമ്പർ	149
കോളേജ്	138
പോലീസ്	127
ഇംഗ്ലീഷ്	126
മെഡിക്കൽ	123

മീറ്റർ	114
റിട്ട	113
സർവീസ്	111
പാർട്ടി	106
അസോസിയേഷൻ	108
കോഴ്സ്	105
ഡയറക്ടർ	105
സെന്റർ	101
റിപ്പോർട്ട്	98
അഡ്വ	98
വൈസ്	98
ആന്റ്	96
ബാങ്ക്	93
കമ്പ്യൂട്ടർ	93
സയൻസ്	91
ഫീസ്	91
സൂപ്പർ	91
പെൻഷൻ	89
ബോർഡ്	84
മാനേജർ	84
സ്കൂൾ	83
ചെയർമാൻ	82
എഞ്ചിനീയറിംഗ്	81

മൊബൈൽ	81
ഹോട്ടൽ	80
സീനിയർ	78
സ്പോർട്സ്	74
ഓഫീസ്	74
കമ്പനി	74
കൗൺസിൽ	66
ഡിപ്ലോമ	65
റാങ്ക്	64
അവാർഡ്	62
ഇലക്ട്രോണിക്സ്	61
എഞ്ചിനീയർ	60
സെന്റർ	60
അക്കാദമി	59
മാനേജ്മെന്റ്	59
ഓഫർ	58
അസിസ്റ്റന്റ്	57
അഡ്മിഷൻ	56
ഓഫീസർ	55
പോയിന്റ്	55
നാഷണൽ	54
ഫെഡറൽ	53
വേൾഡ്	53

ടെൻഡർ	52
ഫുട്ബോൾ	52
സിവിൽ	51
എയർ	51
ഫോട്ടോ	51

#### 4 ചുരുക്കം

ബുലിയൻ ലോജിക് അടിസ്ഥാനമാക്കി പ്രവർത്തിക്കുന്ന ഇലക്ട്രോണിക് കമ്പ്യൂട്ടറുകളിലൂടെ യാന്ത്രികവൽക്കരിക്കാൻ കഴിയുന്ന ഭൂരിഭാഗം മേഖലകൾക്കു വേണ്ടിയും കമ്പ്യൂട്ടർ അധിഷ്ഠിതസംവിധാനങ്ങൾ വികസിപ്പിച്ചെടുക്കുന്ന കാര്യത്തിൽ സാങ്കേതികവിദഗ്ധർ വിജയം കൈവരിച്ചു കഴിഞ്ഞു. എങ്കിലും ഭാഷാപരമായ ദൗത്യങ്ങൾ നിറവേറ്റുന്ന കമ്പ്യൂട്ടർ സാങ്കേതികവിദ്യകൾ ഇനിയും പൂർണ്ണതയിലെത്തിയിട്ടില്ല. കാരണം മനുഷ്യഭാഷയുടെ സങ്കീർണ്ണത തന്നെ. ഭാഷയുടെ ഘടനാപരമായ വിശകലനം എത്രമാത്രം പൂർണ്ണതയിലെത്തുന്നു എന്നതിനെ ആശ്രയിച്ചിരിക്കും ഈ രംഗങ്ങളിലെ പുരോഗതി. അതുകൊണ്ടുതന്നെ ഭാഷയുടെ ഘടന വെളിവാക്കുന്ന പഠനങ്ങൾ ഭാഷാധിഷ്ഠിത സാങ്കേതികവിദ്യയുടെ വളർച്ചയ്ക്ക് സഹായകമാണ്. ഭാഷാപരമായ അറിവുകൾ, കമ്പ്യൂട്ടറിലേക്ക് വിവരങ്ങൾ പ്രവേശിപ്പിക്കുവാനുപയോഗിക്കുന്ന ടെക്സ്റ്റ് എഡിറ്റർ മുതൽ യാന്ത്രിക തർജ്ജമയ്ക്ക് (Machine translation) വേണ്ടി വരെയുള്ള സോഫ്റ്റ്‌വെയറുകളുടെ കാര്യക്ഷമമായ നിർമ്മാണത്തിന് അനിവാര്യമാണ്.

ലിഖിതരൂപത്തിലുള്ള മലയാളഭാഷയിലെ അക്ഷരങ്ങളുടെ ആവൃത്തി നിർണ്ണയം എന്ന പദ്ധതി കേവലം ഭാഷാപരമായ അറിവ് മാത്രം ലക്ഷ്യം വച്ചുള്ളതല്ല. മറിച്ച് മലയാളഭാഷയിലധിഷ്ഠിതമായി പ്രവർത്തിക്കുന്ന കമ്പ്യൂട്ടർ സംവിധാനങ്ങളുടെ നിർമ്മാണത്തിന് അത്യന്താപേക്ഷിതമായ വിവരങ്ങൾ നൽകുക എന്ന പ്രാഥമികലക്ഷ്യമാണ് ഈ പ്രോജക്ടിനുള്ളത്. അക്ഷരങ്ങൾ, വാക്കുകൾ എന്നിവയുടെ ആവൃത്തി നിർണ്ണയം പ്രാഥമികനടപടി മാത്രമേ ആകുന്നുള്ളൂ. വ്യാകരണപരമായും ലെക്സിക്കലായുമുള്ള ഘടന വിശദമാക്കുന്ന പഠനങ്ങൾ ഇനിയും നടക്കേണ്ടിയിരിക്കുന്നു.

## 5. അനുബന്ധം

### ഡാറ്റയും കമ്പ്യൂട്ടർ പ്രോഗ്രാമുകളും അടങ്ങിയ സി.ഡി.

പത്ത് ടെക്സ്റ്റ് ഫയലുകളിലായി ശേഖരിക്കപ്പെട്ടിരിക്കുന്ന ഡാറ്റയിൽ കമ്പ്യൂട്ടർ പ്രോഗ്രാമുകളുപയോഗിച്ചാണ് പരിശോധന നടത്തിയത്. ആവശ്യമായ വിവരങ്ങൾ ലഭിക്കുക എന്ന ഉദ്ദേശത്തോട് കൂടി മാത്രമാണ് പ്രോഗ്രാമുകൾ തയ്യാറാക്കിയിരിക്കുന്നത്. പ്രോഗ്രാമുകളും ഡാറ്റയും ഇതോടൊപ്പം ചേർത്തിരിക്കുന്ന സി.ഡിയിൽ ഉൾക്കൊള്ളിച്ചിട്ടുണ്ട്.